



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : C12N 15/12, 15/11, C07K 14/47, C12Q 1/68, C07K 19/00, 16/18, A61K 38/17, A01K 67/027	A1	(11) International Publication Number: <b>WO 96/25493</b> (43) International Publication Date: 22 August 1996 (22.08.96)
---	----	---

(21) International Application Number: PCT/EP96/00716 (22) International Filing Date: 19 February 1996 (19.02.96) (30) Priority Data: 95200390.3 17 February 1995 (17.02.95) EP (34) Countries for which the regional or international application was filed: NL et al. 95201951.1 14 July 1995 (14.07.95) EP (34) Countries for which the regional or international application was filed: NL et al. (71) Applicant (for all designated States except US): K.U. LEUVEN RESEARCH & DEVELOPMENT (BE/BE); Groot Begijn- hof, Benedenstraat 59, B-3000 Leuven (BE). (71)(72) Applicant and Inventor: BULLERDIEK, Jörn (DE/DE); Weißdornpfad 14, D-28355 Bremen (DE). (72) Inventors; and (75) Inventors/Applicants (for US only): VAN DE VEN, Willem, Jan, Marie (BE/BE); Kasteelstraat 19, B-3051 Oud-Heverlee (BE). SCHOENMAKERS, Henricus, Franciscus, Petrus, Maria (BE/BE); St. Jacobsplein 21/17, B-3000 Leuven	(BE) MOLS, Rafaël (BE/BE); Berthoutlaan 1, B-2220 Hallaar (BE). (74) Agent: VAN SOMEREN, Petronella, Francisca, Hendrika, Maria; Arnold & Siedsma, Sweelinckplein 1, NL-2517 GK The Hague (NL). (81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.
--	--

(54) Title: MULTIPLE-TUMOR ABERRANT GROWTH GENES

## (57) Abstract

The present invention relates to the multi-tumor Aberrant Growth (MAG) gene having the nucleotide sequence of any one of the strands of any one of the members of the High Mobility Group protein genes or LIM protein genes, including modified versions thereof. The gene and its derivatives may be used in various diagnostic and therapeutic applications.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

**MULTIPLE-TUMOR ABERRANT GROWTH GENES**

The present invention relates to the identification of the High Mobility Group (HMG) protein gene family as a family of genes frequently associated with aberrant cell growth as found in a variety of both benign and malignant tumors. The invention in particular relates to the identification of a member of the HMG gene family as the broadly acting chromosome 12 breakpoint region gene involved in a number of tumors, including but not limited to the mesenchymal tumors hamartomas (e.g. breast and lung), lipomas, pleomorphic salivary gland adenomas, uterine leiomyomas, angiomyxomas, fibroadenomas of the breast, polyps of the endometrium, atherosclerotic plaques, and other benign tumors as well as various malignant tumors, including but not limited to sarcomas (e.g. rhabdomyosarcoma, osteosarcoma) and carcinomas (e.g. of breast, lung, skin, thyroid), as well as leukemias and lymphomas. The invention also relates to another member of the HMG gene family that was found to be implicated in breaks in chromosome 6.

Furthermore, the invention concerns the identification of members of the LIM protein family as another type of tumor-type specific breakpoint region genes and frequent fusion partners of the HMG genes in these tumors. The LPP (Lipoma-Preferred Partner) gene of this family is found to be specific for lipomas. The invention relates in particular to the use of the members of the HMG and LIM gene family and their derivatives in diagnosis and therapy.

Multiple independent cytogenetic studies have firmly implicated region q13-q15 of chromosome 12 in a variety of benign and malignant solid tumor types. Among benign solid tumors, involvement of 12q13-q15 is frequently observed in benign adipose tissue tumors [1], uterine leiomyomas [2, 3], and pleomorphic adenomas of the salivary glands [4, 5]. Involvement of the same region has also been reported for endometrial polyps [6, 7] for hemangio-

pericytoma [8], and for chondromatous tumors [9, 10, 11, 12]. Recently, the involvement of chromosome 12q13-q15 was reported in pulmonary chondroid hamartoma [13, 14]. Finally, several case reports of solid tumors with involvement of chromosome region 12q13-q15 have been published; e.g. tumors of the breast [15, 16], diffuse astrocytomas [17], and a giant-cell tumor of the bone [18]. Malignant tumor types with recurrent aberrations in 12q13-q15 include myxoid liposarcoma [19], soft tissue clear-cell sarcoma [20, 21, 22], and a subgroup of rhabdomyosarcoma [23].

Although these studies indicated that the same cytogenetic region of chromosome 12 is often involved in chromosome aberrations, like translocations, in these solid tumors, the precise nature of the chromosome 12 breakpoints in the various tumors is still not known. Neither was it established which genes are affected directly by the translocations.

In previous physical mapping studies [39], the chromosome 12q breakpoints in lipoma, pleomorphic salivary gland adenoma, and uterine leiomyoma were mapped between locus D12S8 and the CHOP gene and it was shown that D12S8 is located distal to CHOP. Recently, it was also found by FISH analysis that the chromosome 12q breakpoints in a hamartoma of the breast, an angiomyxoma and multiple pulmonary chondroid hamartomas are mapping within this DNA interval. In an effort to molecularly clone the genes affected by the chromosome 12q13-q15 aberrations in the various tumors, the present inventors chose directional chromosome walking as a structural approach to define the DNA region encompassing these breakpoints.

As a starting point for chromosome walking, locus D12S8 was used. During these walking studies, it was shown that the chromosomal breakpoints as present in a number of uterine leiomyoma-derived cell lines are clustered within a 445 kb chromosomal segment which has been designated Uterine Leiomyoma Cluster Region on chromosome 12 (ULCR12) [24]. Subsequently, it was found that a 1.7 Mb region on chromosome 12 encompasses the chromosome 12 breakpoints of

uterine leiomyoma-, lipoma-, and salivary gland adenoma-cells, with the breakpoint cluster regions of the various tumor types overlapping [25, "ANNEX 1"]. This 1.7 Mb region on the long arm of chromosome 12, which contains ULCR12  
5 obviously, was designated Multiple Aberration Region (MAR) to reflect this feature. In a regional fine mapping study, MAR was recently assigned to 12q15.

It has thus been found that essentially all breakpoints of chromosome 12 map in a 1.7 Mb region  
10 indicated herein as the "Multiple Aberration Region" or MAR. Further research revealed that in this region a member of the High Mobility Group gene family, the HMGI-C gene, can be identified as a postulated multi-tumor aberrant growth gene (MAG). The same applies to members of the LIM family which  
15 are also found to be involved in chromosome aberrations. Of these the chromosome 3-derived Lipoma-Preferred Partner (LPP) gene is particularly implicated in lipomas.

LIM proteins are proteins carrying cystein-rich zinc-binding domains, so-called LIM domains. They are  
20 involved in protein-protein interactions [for a review see ref. 80]. One of the members of the protein family is the now disclosed LPP protein mapping at chromosome 3.

According to the present invention the aberrations in the HMGI-C gene on chromosome 12 and the LPP gene on  
25 chromosome 3 have been used as a model to reveal the more general concept of the involvement of members of the HMG and LIM gene families in both benign and malignant tumors. To demonstrate that there exists a more general concept of gene families, which, when affected by chromosome rearrangements,  
30 lead to a particular group of tumor growth, the present inventors demonstrated that the HMGI(Y) gene, which is a member of the HMG family, is involved in breaks in chromosome 6.

Although both the HMG and LIM gene families are  
35 known per se, up till the present invention the correlation between these families and tumor inducing chromosome aberrations, like translocations, deletions, insertions and inversions, has not been anticipated. Furthermore, until now

it was not previously demonstrated that alterations in the physiological expression level of the members of the gene family are probably also implicated in tumor development. According to the invention it was demonstrated that in  
5 normal adult cells the expression level of the HMGI-C gene is practically undetectable, whereas in aberrantly growing cells the expression level is significantly increased.

Based on these insights the present invention now provides for the members of the gene families or derivatives  
10 thereof in isolated form and their use in diagnostic and therapeutic applications. Furthermore the knowledge on the location and nucleotide sequence of the genes may be used to study their rearrangements or expression and to identify a possible increase or decrease in their expression level and  
15 the effects thereof on cell growth. Based on this information diagnostic tests or therapeutic treatments may be designed.

In this application the term "Multi-tumor Aberrant Growth (or MAG) gene" will be used to indicate the  
20 involvement of these types of genes in various types of cancer. The term refers to all members of the HMG and LIM gene families involved in non-physiological proliferative growth, and in particular involved in malignant or benign tumors, including atherosclerotic plaques. However,  
25 according to the invention it was furthermore found that even breaks outside the actual gene but in the vicinity thereof might result in aberrant growth. The term MAG gene is therefore also intended to include the immediate vicinity of the gene. The skilled person will understand that the  
30 "immediate vicinity" should be understood to include the surroundings of the gene in which breaks or rearrangements will result in the above defined non-physiological proliferative growth.

The term "wildtype cell" is used to indicate the  
35 cell not harbouring an aberrant chromosome or to a cell having a physiological expression level of the relevant gene. "Wildtype" or "normal" chromosome refers to a non-aberrant chromosome.

The present invention provides for various diagnostic and therapeutic applications that are based on the information that may be derived from the genes. This information not only encompasses its nucleotide sequence or the amino acid sequence of the gene product derived from the gene, but also involves the levels of transcription or translation of the gene.

The invention is thus two-fold. On the one hand the aberration in cell growth may be directly or indirectly caused by the physical breaks that occur in the gene or its vicinity. On the other hand the aberration in cell growth may be caused by a non-physiological expression level of the gene. This non-physiological expression level may be caused by the break, or may be due to another stimulus that activates or deactivates the gene. At present the exact mechanism or origin of the aberrant cell growth is not yet unraveled. However, exact knowledge on this mechanism is not necessary to define methods of diagnosis or treatment.

Diagnostic methods according to the invention are thus based on the fact that an aberration in a chromosome results in a detectable alteration in the chromosomes' appearance or biochemical behaviour. A translocation, for example will result in a first part of the chromosome (and consequently of the MAG gene) having been substituted for another (second) part (further referred to as "first and second substitution parts"). The first part will often appear someplace else on another chromosome from which the second part originates. As a consequence hybrids will be formed between the remaining parts of both (or in cases of triple translocations, even more) chromosomes and the substitution parts provided by their translocation partners. Since it has now been found that the breaks occur in a MAG gene this will result in hybrid gene products of that MAG gene. Markers, such as hybridising molecules like RNA, DNA or DNA/RNA hybrids, or antibodies will be able to detect such hybrids, both on the DNA level, and on the RNA or protein level.

For example, the transcript of a hybrid will still comprise the region provided by the remaining part of the gene/chromosome but will miss the region provided by the substitution part that has been translocated. In the case of  
5 inversions, deletions and insertions the gene may be equally afflicted.

Translocations are usually also cytogenetically detectable. The other aberrations are more difficult to find because they are often not visible on a cytogenetical level.  
10 The invention now provides possibilities for diagnosing all these types of chromosomal aberrations.

In translocations markers or probes based on the MAG gene for the remaining and substitution parts of a chromosome in situ detect the remaining part on the original  
15 chromosome but the substitution part on another, the translocation partner.

In the case of inversions for example, two probes will hybridise at a specific distance in the wildtype gene. This distance might however change due to an inversion. In  
20 situ such inversion may thus be visualized by labeling a set of suitable probes with the same or different detectable markers, such as fluorescent labels. Deletions and insertions may be detected in a similar manner.

According to the invention the above in situ  
25 applications can very advantageously be performed by using FISH techniques. The markers are e.g. two cosmids one of which comprises exons 1 to 3 of the MAG gene, while the other comprises exons 4 and 5. Both cosmids are labeled with different fluorescent markers, e.g. blue and yellow. The  
30 normal chromosome will show a combination of both labels, thus giving a green signal, while the translocation is visible as a blue signal on the remaining part of one chromosome (e.g. 12) while the yellow signal is found on another chromosome comprising the substitution part. In case  
35 the same labels are used for both probes, the intensity of the signal on the normal chromosome will be 100%, while the signal on the aberrant chromosomes is 50%. In the case of

inversions one of the signals shifts from one place on the normal chromosome to another on the aberrant one.

In the above applications a reference must be included for comparison. Usually only one of the two  
5 chromosomes is afflicted. It will thus be very convenient to use the normal chromosome as an internal reference.  
Furthermore it is important to select one of the markers on the remaining or unchanging part of the chromosome and the other on the substitution or inverted part. In the case of  
10 the MAG gene of chromosome 12, breaks are usually found in the big intron between exons 3 and 4 as is shown by the present invention. Furthermore breaks were detected between exons 4 and 5. Probes based on exons 1 to 3 and 4 and 5, or probes based on either exon 4 or on exon 5 are thus very  
15 useful. As an alternative a combination of probes based on both translocation or fusion partners may be used. For example, for the identification of lipomas one may use probes based on exons 1 to 3 of the HMGI-C gene on the one hand and based on the LIM domains of the LPP gene on the  
20 other hand.

Furthermore it was found that breaks might also occur outside the gene, i.e. 5' or 3' thereof. The choice of the probes will then of course include at least one probe hybridising to a DNA sequence located 5' or 3' of the gene.

25 "Probes" as used herein should be widely interpreted and include but are not limited to linear DNA or RNA strands, Yeast Artificial Chromosomes (YACs), or circular DNA forms, such as plasmids, phages, cosmids etc..

These in situ methods may be used on metaphase and  
30 interphase chromosomes.

Besides the above described in situ methods various diagnostic techniques may be performed on a more biochemical level, for example based on alterations in the DNA, RNA or protein, or on changes in the physiological  
35 expression level of the gene.

Basis for the methods that are based on alterations in the chromosome's biochemical behaviour is the fact that by choosing suitable probes, variations in the

length or composition in the gene, transcript or protein may be detected on a gel or blot. Variations in length are visible because the normal gene, transcript(s) or protein(s) will appear in another place on the gel or blot than the  
5 aberrant one(s). In case of a translocation more than the normal number of spots will appear.

Based on the above principle the invention may thus for example relate to a method of diagnosing cells having a non-physiological proliferative capacity,  
10 comprising the steps of taking a biopsy of the cells to be diagnosed, isolating a suitable MAG gene-related macromolecule therefrom, and analysing the macromolecule thus obtained by comparison with a reference molecule originating from cells not showing a non-physiological  
15 proliferative capacity, preferably from the same individual. The MAG gene-related macromolecule may thus be a DNA, an RNA or a protein. The MAG gene may be either a member of the HMG family or of the LIM family.

In a specific embodiment of this type of  
20 diagnostic method the invention comprises the steps of taking a biopsy of the cells to be diagnosed, extracting total RNA thereof, preparing a first strand cDNA of the mRNA species in the total RNA extract or poly-A-selected fraction(s) thereof, which cDNA comprises a suitable tail;  
25 performing a PCR using a MAG gene specific primer and a tail-specific primer in order to amplify MAG gene specific cDNA's; separating the PCR products on a gel to obtain a pattern of bands; evaluating the presence of aberrant bands by comparison to wildtype bands, preferably originating from  
30 the same individual.

As an alternative amplification may be performed by means of the Nucleic Acid Sequence-Based Amplification (NASBA) technique [81] or variations thereof.

In another embodiment the method comprises the  
35 steps of taking a biopsy of the cells to be diagnosed, isolating total protein therefrom, separating the total protein on a gel to obtain essentially individual bands, optionally transferring the bands to a Western blot,

hybridising the bands thus obtained with antibodies directed against a part of the protein encoded by the remaining part of the MAG gene and against a part of the protein encoded by the substitution part of the MAG gene; visualising the  
5 antigen-antibody reactions and establishing the presence of aberrant bands by comparison with bands from wildtype proteins, preferably originating from the same individual.

In a further embodiment the method comprises taking a biopsy of the cells to be diagnosed; isolating  
10 total DNA therefrom; digesting the DNA with one or more so-called "rare cutter" restriction enzymes (typically "6- or more cutters"); separating the digest thus prepared on a gel to obtain a separation pattern; optionally transferring the separation pattern to a Southern blot; hybridising the  
15 separation pattern in the gel or on the blot with a set of probes under hybridising conditions; visualising the hybridisations and establishing the presence of aberrant bands by comparison to wildtype bands, preferably originating from the same individual.

20 Changes in the expression level of the gene may be detected by measuring mRNA levels or protein levels by means of a suitable probe.

Diagnostic methods based on abnormal expression levels of the gene may comprise the steps of taking a sample  
25 of the cells to be diagnosed; isolating mRNA therefrom; and establishing the presence and/or the (relative) quantity of mRNA transcribed from the MAG gene of interest in comparison to a control. Establishing the presence or (relative) quantity of the mRNA may be achieved by amplifying at least  
30 part of the mRNA of the MAG gene by means of RT-PCR or similar amplification techniques. In an alternative embodiment the expression level may be established by determination of the presence or the amount of the gene product (e.g. protein) by means of for example monoclonal  
35 antibodies.

The diagnostic methods of the invention may be used for diseases wherein cells having a non-physiological proliferative capacity are selected from the group

consisting of benign tumors, such as the mesenchymal tumors hamartomas (e.g. breast and lung), adipose tissue tumors (e.g. lipomas), pleomorphic salivary gland adenomas, uterine leiomyomas, angiomyxomas, fibroadenomas of the breast, 5 polyps of the endometrium, atherosclerotic plaques, and other benign tumors as well as various malignant tumors, including but not limited to sarcomas (e.g. rhabdomyosarcoma, osteosarcoma) and carcinomas (e.g. of breast, lung, skin, thyroid). The invention is not limited 10 to the diagnosis and treatment of so-called benign and malignant solid tumors, but the principles thereof have been found to also apply to haematological malignancies like leukemias and lymphomas.

Recent publications indicate that atherosclerotic 15 plaques also involve abnormal proliferation [26] of mainly smooth muscle cells and it was postulated that atherosclerotic plaques constitute benign tumors [27]. Therefore, this type of disorder is also to be understood as a possible indication for the use of the MAG gene family, in 20 particular in diagnostic and therapeutic applications.

As already indicated above it has been found that in certain malignant tumors the expression level of the HMG genes is increased [28]. Until now the relevance of this observation was not understood. Another aspect of the 25 invention thus relates to the implementation of the identification of the MAG genes in therapy. The invention for example provides anti-sense molecules or expression inhibitors of the MAG gene for use in the treatment of diseases involving cells having a non-physiological 30 proliferative capacity by modulating the expression of the gene. A non-physiological high expression may thus be normalised by means of antisense RNA that is either administered to the cell or expressed thereby and binds to the mRNA, or antibodies directed to the gene product, which 35 in turn may result in a normalisation of the cell growth. The examples will demonstrate that expression of antisense RNA in leukemic cells results in a re-differentiation of the cells back to normal.

The invention thus provides derivatives of the MAG gene and/or its immediate environment for use in diagnosis and the preparation of therapeutical compositions, wherein the derivatives are selected from the group consisting of  
5 sense and anti-sense cDNA or fragments thereof, transcripts of the gene or fragments thereof, antisense RNA, triple helix inducing molecule or other types of "transcription clamps", fragments of the gene or its complementary strand, proteins encoded by the gene or fragments thereof, protein  
10 nucleic acids (PNA), antibodies directed to the gene, the cDNA, the transcript, the protein or the fragments thereof, as well as antibody fragments. Besides the use of direct derivatives of the genes and their surroundings (flanking sequences) in diagnosis and therapy, other molecules, like  
15 expression inhibitors or expression enhancers, may be used for therapeutic treatment according to the invention. An example of this type of molecule are ribozymes that destroy RNA molecules.

Besides the above described therapeutic and  
20 diagnostic methods the principles of the invention may also be used for producing a transgenic animal model for testing pharmaceuticals for treatment of MAG gene related malignant or benign tumors and atherosclerotic plaques. One of the examples describes the production of such an animal model.

25 It is to be understood that the principles of the present invention are described herein for illustration purposes only with reference to the HMGI-C gene mapping at chromosome 12 and the HMGI(Y) gene mapping at chromosome 6 and the LPP gene on chromosome 3. Based on the information  
30 provided in this application the skilled person will be able to isolate and sequence corresponding genes of the gene family and apply the principles of this invention by using the gene and its sequence without departing from the scope of the general concept of this invention.

35 The present invention will thus be further elucidated by the following examples which are in no way intended to limit the scope thereof.

## EXAMPLES

### EXAMPLE 1

#### 1. Introduction

This example describes the isolation and analysis  
5 of 75 overlapping YAC clones and the establishment of a YAC  
contig (set of overlapping clones), which spans about 6 Mb  
of genomic DNA around locus D12S8 and includes MAR. The  
orientation of the YAC contig on the long arm of chromosome  
12 was determined by double-color FISH analysis. On the  
10 basis of STS-content mapping and restriction enzyme  
analysis, a long range physical map of this 6 Mb DNA region  
was established. The contig represents a useful resource for  
cDNA capture aimed at identifying genes located in 12q15,  
including the one directly affected by the various  
15 chromosome 12 aberrations.

#### 2. Materials and methods

##### 2.1. Cell lines

Cell lines PK89-12 and LIS-3/SV40/A9-B4 were used  
20 for Chromosome Assignment using Somatic cell Hybrids (CASH)  
experiments. PK89-12, which contains chromosome 12 as the  
sole human chromosome in a hamster genetic background, has  
been described before [29]. PK89-12 cells were grown in DME-  
F12 medium supplemented with 10% fetal bovine serum, 200  
25 IU/ml penicillin, and 200 µg/ml streptomycin. Somatic cell  
hybrid LIS-3/SV40/A9-B4 was obtained upon fusion of myxoid  
liposarcoma cell line LIS-3/SV40, which carries a  
t(12;16)(q13;p11.2), and mouse A9 cells and was previously  
shown to contain der(16), but neither der(12) nor the normal  
30 chromosome 12 [30]. LIS-3/SV40/A9-B4 cells were grown in  
selective AOA-medium (AOA-medium which consisted of DME-F12  
medium supplemented with 10% fetal bovine serum, 0.05 mM  
adenine, 0.05 mM ouabain, and 0.01 mM azaserine). Both cell  
lines were frequently assayed by standard cytogenetic  
35 techniques.

## 2.2. Nucleotide sequence analysis and oligonucleotides.

Nucleotide sequences were determined according to the dideoxy chain termination method using a T7 polymerase sequencing kit (Pharmacia/LKB) or a dsDNA Cycle Sequencing System (GIBCO/BRL). DNA fragments were subcloned in pGEM-3Zf(+) and sequenced using FITC-labelled standard SP6 or T7 primers, or specific primers synthesized based upon newly obtained sequences. Sequencing results were obtained using an Automated Laser Fluorescent (A.L.F.) DNA sequencer (Pharmacia Biotech) and standard 30 cm, 6% Hydrolink<sup>R</sup>, Long Range<sup>TM</sup> gels (AT Biochem). The nucleotide sequences were analyzed using the sequence analysis software Genepro (Riverside Scientific), PC/Gene (IntelliGenetics), the IntelliGenetics Suite software package (IntelliGenetics, Inc.), and Oligo [31]. All oligonucleotides were purchased from Pharmacia Biotech.

## 2.3. Chromosome preparations and fluorescence in situ hybridization (FISH)

FISH analysis of YAC clones was performed to establish their chromosomal positions and to identify chimeric clones. FISH analysis of cosmid clones corresponding to STSs of YAC insert ends were performed to establish their chromosomal positions. Cosmids were isolated from human genomic library CMLW-25383 [32] or the arrayed chromosome 12-specific library constructed at Lawrence Livermore National Laboratory (LL12NC01, ref. 33) according to standard procedures [34]. Routine FISH analysis was performed essentially as described before [30, 35]. DNA was labelled with biotin-11-dUTP (Boehringer) using the protocol of Kievits et al. [36]. Antifade medium, consisting of DABCO (2 g/100 ml, Sigma), 0.1 M Tris-HCL pH 8, 0.02% Thimerosal, and glycerol (90%), and containing propidium iodide (0.5 µg/ml, Sigma) as a counterstain, was added 15 min before specimens were analyzed on a Zeiss Axiophot fluorescence microscope using a double band-pass filter for FITC/Texas red (Omega Optical, Inc.). Results were recorded on Scotch (3M) 640 ASA film.

For the double colour FISH experiments, LLNL12NC01-96C11 was labelled with digoxigenin-11-dUTP (Boehringer) and cosmids LLNL12NC01-1F6 and -193F10, with biotin-11-dUTP. Equal amounts of each probe were combined and this mixture was used for hybridization. After hybridization, slides were incubated for 20 min with Avidin-FITC and then washed as described by Kievits et al. [36]. Subsequent series of incubations in TNB buffer (0.1 M Tris-HCl pH 7.5, 0.15 M NaCl, 0.5% Boehringer blocking agent (Boehringer)) and washing steps were performed in TNT buffer (0.1 M Tris-HCl pH 7.5, 0.15 M NaCl, 0.05% Tween-20); all incubations were performed at 37 °C for 30 min. During the second incubation, Goat- $\alpha$ -Avidin-biotin (Vector) and Mouse- $\alpha$ -digoxigenin (Sigma) were applied simultaneously. During the third incubation, Avidin-FITC and Rabbit- $\alpha$ -Mouse-TRITC (Sigma) were applied. During the last incubation, Goat- $\alpha$ -Rabbit-TRITC (Sigma) was applied. After a last wash in TNT buffer, samples were washed twice in 1 x PBS and then dehydrated through an ethanol series (70%, 90%, 100%). Antifade medium containing 75 ng/ $\mu$ l DAPI (Serva) as counterstain was used. Specimens were analyzed on a Zeiss Axiophot fluorescence microscope as described above.

#### 2.4. Screening of YAC libraries.

YAC clones were isolated from CEPH human genomic YAC libraries mark 1 and 3 [37, 38] made available to us by the Centre d'Étude du Polyphormisme Humain (CEPH). Screening was carried out as previously described [39]. Contaminating Candida parapsylosis, which was sometimes encountered, was eradicated by adding terbinafin (kindly supplied by Dr. Dieter Römer, Sandoz Pharma LTD, Basle, Switzerland) to the growth medium (final concentration: 25  $\mu$ g/ml). The isolated YAC clones were characterized by STS-content mapping, contour-clamped homogeneous electric field (CHEF) gel electrophoresis [40], restriction mapping, and hybridization- and FISH analysis.

## 2.5. PCR reactions

PCR amplification was carried out using a Pharmacia/LKB Gene ATAQ Controller (Pharmacia/LKB) in final volumes of 100  $\mu$ l containing 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.01% gelatine, 2 mM dNTP's, 20 pmole of each amplimer, 2.5 units of Amplitaq (Perkin-Elmer Cetus), and 100 ng (for superpools) or 20 ng (for pools) of DNA. After initial denaturation for 5 min at 94 °C, 35 amplification cycles were performed each consisting of denaturation for 1 min at 94 °C, annealing for 1 min at the appropriate temperature (see Table I) and extension for 1 min at 72 °C. The PCR reaction was completed by a final extension at 72 °C for 5 min. Results were evaluated by analysis of 10  $\mu$ l of the reaction product on polyacrylamide minigels.

## 2.6. Pulsed-field gel electrophoresis and Southern blot analysis

Pulsed-field gel electrophoresis and Southern blot analysis were performed exactly as described by Schoenmakers et al. [39]. Agarose plugs containing high-molecular weight YAC DNA (equivalent to about  $1 \times 10^8$  yeast cells) were twice equilibrated in approximately 25 ml TE buffer (pH 8.0) for 30 min at 50 °C followed by two similar rounds of equilibration at room temperature. Plugs were subsequently transferred to round-bottom 2 ml eppendorf tubes and equilibrated two times for 30 min in 500  $\mu$ l of the appropriate 1 x restriction-buffer at the appropriate restriction temperature. Thereafter, DNA was digested in the plugs according to the suppliers (Boehringer) instructions for 4 h using 30 units of restriction endonuclease per digestion reaction. After digestion, plugs along with appropriate molecular weight markers were loaded onto a 1% agarose / 0.25 x TBE gel, sealed with LMP-agarose and size fractionated on a CHEF apparatus (Biorad) for 18 h at 6.0 V/cm using a pulse angle of 120 degrees and constant pulse times varying from 10 sec (separation up to 300 kbp) to 20 sec (separation up to 500 kbp). In the case of large

restriction fragments, additional runs were performed, aiming at the separation of fragments with sizes above 500 kbp. Electrophoresis was performed at 14 °C in 0.25 x TBE. As molecular weight markers, lambda ladders (Promega) and home-made plugs containing lambda DNA cut with restriction endonuclease HindIII were used. After electrophoresis, gels were stained with ethidium bromide, photographed, and UV irradiated using a stratalinker (Stratagene) set at 120 mJ. DNA was subsequently blotted onto Hybond N<sup>+</sup> membranes (Amersham) for 4-16 h using 0.4 N NaOH as transfer buffer. After blotting, the membranes were dried for 15 min at 80 °C, briefly neutralised in 2 x SSPE, and prehybridised for at least 3 h at 42 °C in 50 ml of a solution consisting of 50% formamide, 5 x SSPE, 5 x Denhardts, 0.1% SDS and 200 µg/ml heparin. Filters were subsequently hybridised for 16 h at 42 °C in 10 ml of a solution consisting of 50% formamide, 5 x SSPE, 1 x Denhardts, 0.1% SDS, 100 µg/ml heparin, 0.5% dextran sulphate and 2-3 x 10<sup>6</sup> cpm/ml of labelled probe. Thereafter, membranes were first washed two times for 5 min in 2 x SSPE/0.1% SDS at room temperature, then for 30 min in 2 x SSPE/0.1% SDS at 42 °C and, finally, in 0.1 x SSPE/0.1% SDS for 20 min at 65 °C. Kodak XAR-5 films were exposed at -80 °C for 3-16 h, depending on probe performance. Intensifying screens (Kyokko special 500) were used.

25

## 2.7. Generation of STSs from YAC insert ends

STSs from YAC insert ends were obtained using a vectorette-PCR procedure in combination with direct DNA sequencing analysis, essentially as described by Geurts et al. [41]. Primer sets were developed and tested on human genomic DNA, basically according to procedures described above. STSs will be referred to throughout this application by their abbreviated names (for instance: RM1 instead of STS 12-RM1) for reasons of convenience.

35

### 3. Results

#### 3.1. Assembly of a YAC Contig around locus D12S8

In previous studies [39], a 800 kb YAC contig around D12S8 was described. This contig consisted of the following three partially overlapping, non-chimeric CEPH YAC clones: 258F11, 320F6, and 234G11. This contig was used as starting point for a chromosome walking project to define the DNA region on the long arm of chromosome 12 that encompasses the breakpoints of a variety of benign solid tumors, which are all located proximal to D12S8 and distal to CHOP. Initially, chromosome walking was performed bidirectionally until the size of the contig allowed reliable determination of the orientation of it. In the bidirectional and subsequent unidirectional chromosome walking steps, the following general procedure was used. First, rescuing and sequencing the ends of YAC clones resulted in DNA markers characterizing the left and right sides of these (Table I). Based on sequence data of the ends of forty YAC inserts, primer sets were developed for specific amplification of DNA; establishing STSs (Table II). Their localization to 12q13-qter was determined by CASH as well as FISH after corresponding cosmid clones were isolated. It should be noted that isolated YAC clones were often evaluated by FISH analysis too, thus not only revealing the chromosomal origin of their inserts but also, for a number of cases, establishing and defining their chimeric nature. Moreover, it should be emphasized that data obtained by restriction endonuclease analysis of overlapping YAC clones were also taken into account in the YAC clone evaluation and subsequent alignment. With the sequentially selected and evaluated primer sets, screening of the YAC and cosmid libraries was performed to isolate the building blocks for contig-assembly. Therefore, contig-assembly was performed using data derived from FISH- and STS-content mapping as well as restriction endonuclease analysis. Using this approach, we established a YAC contig consisting of 75 overlapping YAC clones, covering approximately 6 Mb of DNA (Fig. 1). This contig appeared to encompass the chromosome

12 breakpoints of all tumor-derived cell lines studied [39]. Characteristics of the YACs that were used to build this contig are given in Table I.

### 5 3.2. Establishment of the Chromosomal Orientation of the YAC Contig

To allow unidirectional chromosome walking towards the centromere of chromosome 12, the orientation of the DNA region flanked by STSs RM14 and RM26 (approximate size: 1450  
10 kb) was determined by double-color interphase FISH analysis. Cosmid clones corresponding to these STSs (i.e. LL12NC01-1F6 (RM14) and LL12NC01-96C11 (RM26)) were differentially labelled to show green or red signals, respectively. As a reference locus, cosmid LL12NC01-193F10 was labelled to show  
15 green signals upon detection. LL12NC01-193F10 had previously been mapped distal to the breakpoint of LIS-3/SV40 (i.e. CHOP) and proximal to the chromosome 12q breakpoints in lipoma cell line Li-14/SV40 and uterine leiomyoma cell line LM-30.1/SV40. LL12NC01-1F6 and LL12NC01-96C11 were found to  
20 be mapping distal to the 12q breakpoints in lipoma cell line Li-14/SV40 and uterine leiomyoma cell line LM-30.1/SV40. Therefore, LL12NC01-193F10 was concluded to be mapping proximal to both RM14 and RM26 (unpublished results). Of 150 informative interphases scored, 18% showed a signal-order of  
25 red-green-green whereas 72% showed a signal order of green-red-green. Based upon these observations, we concluded that RM26 mapped proximal to RM14, and therefore we continued to extend the YAC contig from the RM26 (i.e. proximal) side of our contig only. Only the cosmids containing RM14 and RM26  
30 were ordered by double-color interphase mapping; the order of all others was deduced from data of the YAC contig. Finally, it should be noted that the chromosomal orientation of the contig as proposed on the basis of results of the double-color interphase FISH studies was independently  
35 confirmed after the YAC contig had been extended across the chromosome 12 breakpoints as present in a variety of tumor cell lines. This confirmatory information was obtained in extensive FISH studies in which the positions of YAC and

cosmid clones were determined relative to the chromosome 12q13-q15 breakpoints of primary lipomas, uterine leiomyomas, pleomorphic salivary gland adenomas, and pulmonary chondroid hamartomas or derivative cell lines [24, 42, 25, 43].

### 3.3. Construction of a Rare-Cutter Physical Map from the 6 Mb YAC Contig around D12S8

Southern blots of total yeast plus YAC DNA, digested to completion with rare-cutter enzymes (see Materials and Methods) and separated on CHEF gels, were hybridized sequentially with i) the STS used for the initial screening of the YAC in question, ii) pYAC4 right arm sequences, iii) pYAC4 left arm sequences, and iv) a human ALU-repeat probe (BLUR-8). The long-range restriction map that was obtained in this way, was completed by probing with PCR-isolated STSs/YAC end probes. Occasionally double-digests were performed.

Restriction maps of individual YAC clones were aligned and a consensus restriction map was established. It is important to note here that the entire consensus rare-cutter map was supported by at least two independent clones showing a full internal consistency.

### 3.4. Physical mapping of CA repeats and monomorphic STSs/ESTs

Based upon integrated mapping data as emerged from the Second International Workshop on Human Chromosome 12 [44], a number of published markers was expected to be mapping within the YAC contig presented here. To allow full integration of our mapping data with those obtained by others, a number of markers were STS content-mapped on our contig, and the ones found positive were subsequently sublocalized by (primer-)hybridization on YAC Southern blots. Among the markers that were found to reside within the contig presented here were CA repeats D12S313 (AFM207xf2) and D12S335 (AFM273vg9) [45], D12S375 (CHLC GATA3F02), and D12S56 [46]. Furthermore, the interferon

gamma gene (IFNG) [47], the ras-related protein gene Rap1B [48], and expressed sequence tag EST01096 [49] were mapped using primer sets which we developed based on publicly available sequence data (see Table II). Markers which were  
5 tested and found negative included D12S80 (AFM102xd6), D12S92 (AFM203va7), D12S329 (AFM249xh9) and D12S344 (AFM296xd9).

#### 4. Discussion

10 In the present example the establishment of a YAC contig and rare-cutter physical map covering approximately 6 Mb on 12q15, a region on the long arm of human chromosome 12 containing MAR in which a number of recurrent chromosomal aberrations of benign solid tumors are known to be mapping  
15 was illustrated.

The extent of overlap between individual YACs has been carefully determined, placing the total length of the contig at approximately 6 Mb (Fig. 1). It should be noted that our sizing-data for some of the YAC clones differ  
20 slightly from the sizes determined by CEPH [50]. It is our belief that this is most probably due to differences in the parameters for running the pulsed-field gels in the different laboratories.

Using restriction mapping and STS-content  
25 analysis, a consensus long range physical map (Fig. 1) was constructed. The entire composite map is supported by at least two-fold coverage. In total over 30 Mb of YAC DNA was characterized by restriction and STS content analysis, corresponding to an average contig coverage of about 5  
30 times. Although the "inborn" limited resolution associated with the technique of pulsed-field electrophoresis does not allow very precise size estimations, comparison to restriction mapping data obtained from a 500 kb cosmid contig contained within the YAC contig presented here showed  
35 a remarkable good correlation. Extrapolating from the cosmid data, we estimate the accuracy of the rare-cutter physical map presented here at about 10 kb.

The results of our physical mapping studies allowed integration of three gene-specific as well as five anonymous markers isolated by others (indicated in between arrows in Fig. 1). The anonymous markers include one monomorphic and four polymorphic markers. Five previously published YAC-end-derived single copy STSs (RM1, RM4, RM5, RM7, and RM21) as well as four published CA repeats (D12S56, D12S313, D12S335, and D12S375) and three published gene-associated STSs/ESTs (RAP1B, EST01096, and IFNG) have been placed on the same physical map and this will facilitate (linkage-) mapping and identification of a number of traits/disease genes that map in the region. Furthermore, we were able to place onto the same physical map, seventy two YAC-end-derived (Table I) and eight cosmid-end- or inter-ALU-derived DNA markers (CH9, RM1, RM110, RM111, RM130, RM131, RM132, and RM133), which were developed during the process of chromosome walking. The PYTHIA automatic mail server at PYTHIA@anl.gov was used to screen the derived sequences of these DNA markers for the presence of repeats. For forty three of these seventy two DNA markers (listed in Table II), primer sets were developed and the corresponding STSs were determined to be single copy by PCR as well as Southern blot analysis of human genomic DNA. The twenty nine remaining DNA markers (depicted in the yellow boxes) represent YAC-end-derived sequences for which we did not develop primer sets. These YAC-end sequences are assumed to be mapping to chromosome 12 on the basis of restriction mapping. The final picture reveals an overall marker density in this region of approximately one within every 70 kb.

The analysis of the contig presented here revealed many CpG-rich regions, potentially HTF islands, which are known to be frequently associated with housekeeping genes. These CpG islands are most probably located at the 5' ends of as yet unidentified genes: it has been shown that in 90% of cases in which three or more rare-cutter restriction sites coincide in YAC DNA there is an associated gene [51]. This is likely to be an underestimate of the number of genes yet to be identified in this region because 60% of tissue-

specific genes are not associated with CpG islands [52] and also because it is possible for two genes to be transcribed in different orientations from a single island [53].

While several of the YAC clones that were isolated from the CEPH YAC library mark 1 were found to be chimeric, overlapping YAC clones that appeared to be non-chimeric based on FISH, restriction mapping and STS content analysis could be obtained in each screening, which is in agreement with the reported complexity of the library. The degree of chimerism for the CEPH YAC library mark 1 was determined at 18% (12 out of 68) for the region under investigation here. The small number of YACs from the CEPH YAC library mark 3 (only 7 MEGA YACs were included in this study) did not allow a reliable estimation of the percentage of chimeric clones present in this library. The average size of YACs derived from the mark 1 library was calculated to be 381 kb ; non-chimeric YACs (n=58) had an average size of 366 kb while chimeric YACs (n=12) were found to have a considerable larger average size; i.e. 454 kb.

In summary, we present a 6 Mb YAC contig corresponding to a human chromosomal region which is frequently rearranged in a variety of benign solid tumors. The contig links over 84 loci, including 3 gene-associated STSs. Moreover, by restriction mapping we have identified at least 12 CpG islands which might be indicative for genes residing there. Finally, four CA repeats have been sublocalized within the contig. The integration of the genetic, physical, and transcriptional maps of the region provides a basic framework for further studies of this region of chromosome 12. Initial studies are likely to focus on MAR and ULCR12, as these regions contain the breakpoint cluster regions of at least three distinct types of solid tumors. The various YAC clones we describe here are valuable resources for such studies. They should facilitate the search for genes residing in this area and the identification of those directly affected by the chromosome 12q aberrations of the various benign solid tumors.

**EXAMPLE 2****1. Introduction**

It was found that the 1.7 Mb Multiple Aberration Region on human chromosome 12q15 harbors recurrent  
5 chromosome 12 breakpoints frequently found in different benign solid tumor types. In this example the identification of an HMG gene within MAR that appears to be of pathogenetical relevance is described. Using a positional cloning approach, the High Mobility Group protein gene HMGI-  
10 C was identified within a 175 kb segment of MAR and its genomic organization characterized. By FISH, within this gene the majority of the breakpoints of seven different benign solid tumor types were pinpointed. By Southern blot and 3'-RACE analysis, consistent rearrangements in HMGI-C  
15 and/or expression of altered HMGI-C transcripts were demonstrated. These results indicate a link between a member of the HMG gene family and benign solid tumor development.

**2. Materials and methods****2.1. Cell culture and primary tumor specimens.**

Tumor cell lines listed in Fig. 3 were established by a transfection procedure [54] and have been described before in [39, 24] and in an article of Van de Ven et al.,  
25 Genes Chromosom. Cancer 12, 296-303 (1995) enclosed with this application as ANNEX 1. Cells were grown in TC199 medium supplemented with 20% fetal bovine serum and were assayed by standard cytogenetic techniques at regular intervals. The human hepatocellular carcinoma cell lines Hep  
30 3B and Hep G2 were obtained from the ATCC (accession numbers ATCC HB 8064 and ATCC HB 8065) and cultured in DMEM/F12 supplemented with 4% Ultrosor (Gibco/BRL). Primary solid tumors were obtained from various University Clinics.

**35 2.2. YAC and cosmid clones**

YAC clones were isolated from the CEPH mark 1 [57] and mark 3 [58] YAC libraries using a combination of PCR-based screening [59] and colony hybridization analysis.

Cosmid clones were isolated from an arrayed human chromosome 12-specific cosmid library (LL12NC01) [60] obtained from Lawrence Livermore National Laboratory (courtesy P. de Jong). LL12NC01-derived cosmid clones are indicated by their 5 microtiter plate addresses; i.e. for instance 27E12.

Cosmid DNA was extracted using standard techniques involving purification over Qiagen tips (Diagen). Agarose plugs containing high-molecular weight yeast + YAC DNA (equivalent to  $1 \times 10^9$  cells  $\text{ml}^{-1}$ ) were prepared as described 10 before [61]. Plugs were thoroughly dialysed against four changes of 25 ml  $T_{10}E_1$  (pH 8.0) followed by two changes of 0.5 ml 1 x restriction buffer before they were subjected to either pulsed-field restriction enzyme mapping or YAC-end rescue. YAC-end rescue was performed using a vectorette-PCR 15 procedure in combination with direct solid phase DNA sequencing, as described before in reference 61. Inter-Alu PCR products were isolated using published oligonucleotides TC65 or 517 [62] to which SalI-tails were added to facilitate cloning. After sequence analysis, primer pairs 20 were developed using the OLIGO computer algorithm [61].

### 2.3. DNA labelling

DNA from YACs, cosmids, PCR products and oligonucleotides was labelled using a variety of techniques. 25 For FISH, cosmid clones or inter-Alu PCR products of YACs were biotinylated with biotin-11-dUTP (Boehringer) by nick translation. For filter hybridizations, probes were radio-labelled with  $\alpha$ - $^{32}\text{P}$ -dCTP using random hexamers [62]. In case of PCR-products smaller than 200 bp in size, a similar 30 protocol was applied, but specific oligonucleotides were used to prime labelling reactions. Oligonucleotides were labelled using  $\gamma$ - $^{32}\text{P}$ -ATP.

### 2.4. Nucleotide sequence analysis and PCR amplification

35 Nucleotide sequences were determined as described in Example 1. Sequencing results were analyzed using an A.L.F. DNA sequencer<sup>TM</sup> (Pharmacia Biotech) on standard 30 cm, 6% Hydrolink<sup>R</sup>, Long Range<sup>TM</sup> gels (AT Biochem). PCR

amplifications were carried out essentially as described before [39].

### 2.5. Rapid amplification of cDNA ends (RACE)

5           Rapid amplification of 3' cDNA-ends (3'-RACE) was performed using a slight modification of part of the GIBCO/BRL 3'-ET protocol. For first strand cDNA synthesis, adapter primer (AP2) AAG GAT CCG TCG ACA TC(T)<sub>17</sub> was used. For both initial and secondary rounds of PCR, the universal  
10 amplification primer (UAP2) CUA CUA CUA CUA AAG GAT CCG TCG ACA TC was used as "reversed primer". In the first PCR round the following specific "forward primers" were used: i) 5'-CTT CAG CCC AGG GAC AAC-3' (exon 1), ii) 5'-CAA GAG GCA GAC CTA GGA-3' (exon 3), or iii) 5'-AAC AAT GCA ACT TTT AAT TAC  
15 TG-3' (3'-UTR). In the second PCR round the following specific forward primers (nested primers as compared to those used in the first round) were used: i) 5'-CAU CAU CAU CAU CGC CTC AGA AGA GAG GAC-3' (exon 1), ii) 5'-CAU CAU CAU CAU GTT CAG AAG AAG CCT GCT-3' (exon 4), or iii) 5'-CAU CAU  
20 CAU CAU TTG ATC TGA TAA GCA AGA GTG GG-3' (3'-UTR). CUA/CAU-tailing of the nested, specific primers allowed the use of the directional CloneAmp cloning system (GIBCO/BRL).

## 3. Results

### 25 3.1. Development of cosmid contig and STS map of MAR segment

During the course of a positional cloning effort focusing on the long arm of human chromosome 12, we constructed a yeast artificial chromosome (YAC) contig spanning about 6 Mb and consisting of 75 overlapping YACs.  
30 For a description thereof reference is made to Example 1. This contig encompasses MAR (see also Fig. 2), in which most of the chromosome 12q13-q15 breakpoints as present in a variety of primary benign solid tumors (34 tumors of eight different types tested sofar; Table 5) and tumor cell lines  
35 (26 tested sofar, derived from lipoma, uterine leiomyoma, and pleomorphic salivary gland adenoma; Fig. 3) appear to cluster. We have developed both a long-range STS and rare cutter physical map of MAR and found, by FISH analysis, most

of the breakpoints mapping within the 445 kb subregion of MAR located between STSs RM33 and RM98 (see Fig. 2 and 3). FISH experiments, including extensive quality control, were performed according to routine procedures as described before [25, 39, 24, 42, 36] To further refine the distribution of breakpoints within this 445 kb MAR segment, a cosmid contig consisting of 54 overlapping cosmid clones has been developed and a dense STS map (Fig. 2) established. The cosmid contig was double-checked by comparison to the rare cutter physical map and by STS content mapping.

### 3.2. Clustering of the chromosome 12q breakpoints within a 175 kb DNA segment of MAR

The chromosome 12q breakpoints in the various tumor cell lines studied was pinpointed within the cosmid contig by FISH (Fig. 3). As part of our quality control FISH experiments [25, 39, 24, 42], selected cosmids were first tested on metaphase spreads derived from normal lymphocytes. FISH results indicated that the majority (at least 18 out of the 26 cases) of the chromosome 12 breakpoints in these tumor cell lines were found to be clustering within the 175 kb DNA interval between RM99 and RM133, indicating this interval to constitute the main breakpoint cluster region. FISH results obtained with Li-501/SV40 indicated that part of MAR was translocated to an apparently normal chromosome 3; a chromosome aberration overseen by applied cytogenetics. Of interest to note, finally, is the fact that the breakpoints of uterine leiomyoma cell lines LM-5.1/SV40, LM-65/SV40, and LM-608/SV40 were found to be mapping within the same cosmid clone; i.e. cosmid 27E12.

We also performed FISH experiments on eight different types of primary benign solid tumors with chromosome 12q13-q15 aberrations (Table 4). A mixture of cosmid clones 27E12 and 142H1 was used as molecular probe. In summary, the results of the FISH studies of primary tumors were consistent with those obtained for the tumor cell lines. The observation that breakpoints of each of the seven different tumor types tested were found within the

same 175 kb DNA interval of MAR suggested that this interval is critically relevant to the development of these tumors and, therefore, might harbor the putative MAG locus or critical part(s) of it.

5

### 3.3. Identification of candidate genes mapping within MAR

In an attempt to identify candidate genes mapping within the 175 kb subregion of MAR between STSs RM99 and RM133, we used 3'-terminal exon trapping and genomic  
10 sequence sampling (GSS) [63]. Using the GSS approach, we obtained DNA sequence data of the termini of a 4.9 kb BamHI subfragment of cosmid 27E12, which was shown by FISH analysis to be split by the chromosome 12 aberrations in three of the uterine leiomyoma cell lines tested. A BLAST  
15 [64] search revealed that part of these sequences displayed sequence identity with a publicly available partial cDNA sequence (EMBL accession # Z31595) of the high mobility group (HMG) protein gene HMGI-C [65], which is a member of the HMG gene family [66]. In light of these observations,  
20 HMGI-C was considered a candidate MAG gene and studied in further detail.

### 3.4. Genomic organization of HMGI-C and rearrangements in benign solid tumors

25 Since only 1200 nucleotides of the HMGI-C transcript (reported size approximately 4 kb [65, 67]) were publicly available, we first determined most of the remaining nucleotide sequences of the HMGI-C transcript (GenBank, # U28749). This allowed us to subsequently  
30 establish the genomic organization of the gene. Of interest to note about the sequence data is that a CT-repeat is present in the 5'-UTR of HMGI-C and a GGGGT-pentanucleotide repeat in the 3'-UTR, which might be of regulatory relevance. Comparison of transcribed to genomic DNA  
35 sequences (GenBank, # U28750, U28751, U28752, U28753, and U28754) of the gene revealed that HMGI-C contains at least 5 exons (Fig. 2). Transcriptional orientation of the gene is directed towards the telomere of the long arm of the

chromosome. Each of the first three exons encode a putative DNA binding domain (DBD), and exon 5 encodes an acidic domain, which is separated from the three DBDs by a spacer domain encoded by exon 4. The three DBD-encoding exons are positioned relatively close together and are separated by a large intron of about 140 kb from the two other exons, which in turn are separated about 11 kb from each other. Of particular interest to emphasize here is that the five exons are dispersed over a genomic region of at least 160 kb, thus almost covering the entire 175 kb main MAR breakpoint cluster region described above. Results of molecular cytogenetic studies, using a mixture of cosmid 142H1 (containing exons 1-3) and 27E12 (containing exons 4 and 5) as molecular probe, clearly demonstrate that the HMGI-C gene is directly affected by the observed chromosome 12 aberrations in the majority of the tumors and tumor cell lines that were evaluated (Fig. 3; Table 4). These cytogenetic observations were independently confirmed by Southern blot analysis in the case of LM-608/SV40 (results not shown) LM-30.1/SV40 [24], and Ad-312/SV40; probes used included CH76, RM118-A, and EM26. The failure to detect the breakpoints of LM-65/SV40, LM-609/SV40, Ad-211/SV40, Ad-263/SV40, Ad-302/SV40, Li-14/SV40, and Li-538/SV40 with any of these three probes was also consistent with the FISH data establishing the relative positions of the breakpoints in MAR (cf. Fig. 3). These results made HMGI-C a prime candidate to be the postulated MAG gene.

### 3.5. Expression of aberrant HMGI-C transcripts in benign solid tumor cells.

In the context of follow-up studies, it was of interest to test for possible aberrant HMGI-C expression. Initial Northern blot studies revealed that transcripts of HMGI-C could not be detected in a variety of normal tissues (brain, heart, lung, liver, kidney, pancreas, placenta, skeletal muscle) tested as well as in a number of the tumor cell lines listed in Fig. 3 (data not shown). It is known that HMGI-C mRNA levels in normal differentiated tissues are

very much lower than in malignant tissues [65, 67]. As a control in our Northern studies, we included hepatoma cell line Hep 3B, which is known to express relatively high levels of HMGI-C. We readily detected two major HMGI-C transcripts, approximately 3.6 and 3.2 kb in size; with the differences in molecular weight most likely due to differences in their 5'-noncoding regions. In an alternative and more sensitive approach to detect HMGI-C or 3'-aberrant HMGI-C transcripts, we have performed 3'-RACE experiments.

10 In control experiments using a number of tissues with varying HMGI-C transcription levels (high levels in Hep 3B hepatoma cells, intermediate in Hep G2 hepatoma cells, and low in myometrium, normal fat tissue, and pseudomyxoma), we obtained 3'-RACE clones which, upon molecular cloning and

15 nucleotide sequence analysis, appeared to represent perfect partial cDNA copies of 3'-HMGI-C mRNA sequences; no matter which of the three selected primer sets was used (see Methodology). RACE products most likely corresponding to cryptic or aberrantly spliced HMGI-C transcripts were

20 occasionally observed; their ectopic sequences were mapped back to HMGI-C intron 3 or 4.

In similar 3'-RACE analysis of ten different primary tumors or tumor cell lines derived from lipoma, uterine leiomyoma, and pleomorphic salivary gland adenoma,

25 we detected both constant and unique PCR products. The constant PCR products appeared to represent, in most cases, perfect partial cDNA copies of 3'-HMGI-C mRNA sequences. They most likely originated from a presumably unaffected HMGI-C allele and might be considered as internal controls.

30 The unique PCR products of the ten tumor cell samples presented here appeared to contain ectopic sequences fused to HMGI-C sequences. In most cases, the ectopic sequences were found to be derived from the established translocation partners, thus providing independent evidence for

35 translocation-induced rearrangements of the HMGI-C gene. Information concerning nucleotide sequences, diversion points, and chromosomal origins of the ectopic sequences of these RACE products is summarized in Table 5. It should be

noted that chromosomal origins of ectopic sequences was established by CASH (Chromosome Assignment using Somatic cell Hybrids) analysis using the NIGMS Human/Rodent Somatic Hybrid Mapping Panel 2 obtained from the Coriell Cell Repositories. Chromosomal assignment was independently confirmed by additional data for cases pCH1111, pCH172, pCH174, pCH193, and pCH117, as further outlined in Table 5. Taking into account the limitations of conventional cytogenetic analysis, especially in cases with complex karyotypes, the chromosome assignments of the ectopic sequences are in good agreement with the previous cytogenetic description of the translocations.

Somewhat unexpected were the data obtained with Ad-312/SV40, as available molecular cytogenetic analysis had indicated its chromosome 12 breakpoint to map far outside the HMGI-C gene; over 1 Mb [42]. The ectopic sequences appeared to originate from chromosome 1 (more precisely from a segment within M.I.T. YAC contig WC-511, which is partially mapping at 1p22), the established translocation partner (Fig. 2). Further molecular analysis is required to precisely define the effect on functional expression of the aberrant HMGI-C gene in this particular case. Of further interest to note here, is that the sequences coming from chromosome 1 apparently remove the GGGGT repeat observed in the 3'-UTR region of HMGI-C, as this repeat is not present in the RACE product. In contrast, in primary uterine leiomyoma LM-#58 (t(8;12)(q24;q14-q15)), which was shown to have its breakpoint also in the 3'-UTR, this repeat appeared to be present in the RACE product. Therefore, removal of this repeat is most probably not critical for tumor development. The results with primary tumor LM-#168.1, in which the X chromosome is the cytogenetically assigned translocation partner, revealed that the ectopic sequences were derived from chromosome 14; the preferential translocation partner in leiomyoma. It is possible that involvement of chromosome 14 cannot be detected by standard karyotyping in this particular case, as turned out to be the case for Li-501/SV40. In primary lipoma Li-#294

(t(8;12)(q22;q14)), two alternative ectopic sequences were detected. Additional CASH analysis using a hybrid cell mapping panel for regional localization of probes to human chromosome 8 [68] revealed that these were both derived from 5 chromosome 8q22-qter (Table 5). It is very well possible that these RACE products correspond to alternatively spliced transcripts. Finally, in four of the cases (Table 5, cases pCH114, pCH110, pCH109, pCH116), the RACE products appeared to correspond to cryptic or aberrantly spliced HMGI-C 10 transcripts, as the corresponding ectopic sequences were found to be derived from either HMGI-C intron 3 or 4. Such RACE products have also been observed in the control experiments described above. In conclusion, the detection of aberrant HMGI-C transcripts in the tumor cells provides 15 additional strong support of HMGI-C being consistently rearranged by the various chromosome 12 aberrations. It should be noted that the aberrant HMGI-C transcripts in the various cases should be characterized in full length before any final conclusion can be drawn about biological 20 implications.

A first and preliminary evaluation of isolated ectopic sequences revealed in phase open reading frames of variable length. In the case of primary tumor LM-#25, for instance, already the second codon in the ectopic sequences 25 appeared to be a stop codon (Table 5). A note of caution is appropriate here, as sequence data have been obtained only for clones that were produced via two rounds of extensive (probably mutations inducing) PCR. For Li-501/SV40, it is of interest to note that, in Northern blot analysis, the 30 isolated ectopic sequences detected a transcript of over 10 kb in a variety of tissues, including heart, kidney, liver, lung, pancreas, placenta, and skeletal muscle, but not in brain (data not shown). As chromosome 3 is the preferred partner in the chromosome 12q13-q15 translocations in 35 lipomas and the chromosome 3 breakpoints of various lipomas were found to be spanned by YAC clone CEPH192B10, the detected transcript might correspond to a putative lipoma-preferred partner gene (LPP).

#### 4. Discussion

In ANNEX 1 it was demonstrated that the chromosome 12q13-q15 breakpoints of lipoma, pleomorphic salivary gland adenoma, and uterine leiomyoma, irrespective of their cytogenetic assignments in the past to segment q13, q14, or q15 of chromosome 12, all cluster within the 1.7 Mb DNA interval designated MAR. In support of the claimed clustering of breakpoints is a recent study by Schoenberg Fejzo et al. [14], identifying a CEPH mega-YAC spanning the chromosome 12 translocation breakpoints in two of the three tumor types. In the present study, we have conclusively demonstrated by FISH analysis that chromosome 12 breakpoints of seven different solid tumor types are clustering within a relatively small (175 kb) segment of MAR. For some tumor cell lines, Southern blot data were obtained and these were always in support of the FISH results. From all these observations, we conclude that this segment of MAR constitutes a major target area for the chromosome 12 aberrations in these tumors and that it is likely to represent the postulated MAG locus: the multi-tumor aberrant growth locus that might be considered as common denominator in these tumors.

Within the 175 kb MAR segment, we have identified the HMGI-C gene and determined characteristics of its genomic organization. Structurally, the HMGI-C-encoded phosphoprotein consists of three putative DNA binding domains, a spacer region, and an acidic carboxy-terminal domain, and contains potential sites of phosphorylation for both casein kinase II and p34/cdc2 [65, 67]. We have provided strong evidence that HMGI-C is a prime candidate target gene involved in the various tumor types studied here. In FISH studies, the breakpoints of 29 out of 33 primary tumors were found to be mapping between two highly informative cosmids 142H1 and 27E12; the first one containing the three DBD-encoding exons and the second one, the remaining exons that code for the two other domains. Therefore, the majority of the breakpoints map within the gene, most of them probably within the 140 kb intron (intron

3), which is also in line with FISH results obtained with the 26 tumor cell lines that were evaluated. It should also be noted that the 5'-end of the HMGI-C gene is not yet fully characterized. As HMGI(Y), another member of this gene family, is known to possess various alternative first exons [69], the size of the HMGI-C gene might be larger than yet assumed. Further support that HMGI-C is affected by the chromosome 12 aberrations can be deduced from the results of the 3'-RACE experiments. Aberrant HMGI-C transcripts were detected in tumor cells, consisting of transcribed HMGI-C sequences fused to newly acquired sequences, in most cases clearly originating from the chromosomes that were cytogenetically identified as the translocation partners. It is noteworthy that many chromosomes have been found as translocation partner in the tumors studied. This observed heterogeneity in the reciprocal breakpoint regions involved in these translocations resembles that of a variety of hematological malignancies with chromosome 11q23 rearrangements involving the MLL gene [70], the translational product of which carries an amino-terminal motif related to the DNA-binding motifs of HMGI proteins.

An intriguing issue pertains to the effect of the chromosome 12 aberrations on expression of the HMGI-C gene and the direct physiological implications. Some functional characteristics of HMGI-C are known or may be deduced speculatively from studies of other family members. As it binds in the minor groove of DNA, it has been suggested that HMGI-C may play a role in organising satellite chromatin or act as a transcription factor [71, 72]. Studies on HMGI(Y), which is the member most closely related to HMGI-C, have suggested that HMGI(Y) may function as a promoter-specific accessory factor for NF- $\kappa$  B transcriptional activity [73]. HMGI(Y) has also been shown to stimulate or inhibit DNA binding of distinct transcriptional factor ATF-2 isoforms [74]. Both studies indicate that the protein may simply constitute a structural component of the transcriptional apparatus functioning in promoter/enhancer contexts. In a recent report on high mobility group protein 1 (HMGI), yet

another member of the HMG gene family with a similar domain structure as HMGI-C and acting as a quasi-transcription factor in gene transcription, a truncated HMGI protein lacking the acidic carboxy-terminal region was shown to  
5 inhibit gene transcription [75]. It was put forward that the acidic terminus of the HMGI molecule is essential for the enhancement of gene expression in addition to elimination of the repression caused by the DNA binding. As most of the chromosome 12 breakpoints seem to occur in the 140 kb  
10 intron, separation of the DBDs from the acidic carboxy-terminal domain seems to occur frequently. In case the acidic domain in HMGI-C has a similar function as the one in HMGI(Y), the result of the chromosome 12 aberrations is likely to affect gene expression. Finally, it should be  
15 noted that the fate of the sequences encoding the acidic carboxy-terminal region is not yet known.

As HMGI-C is the first member of the HMG gene family that might be implicated in the development of benign tumors, the question arises as to whether other members of  
20 this family could also be involved. The HMG protein family consists of three subfamilies: i) the HMGI and 2 type proteins, which have been found to enhance transcription in vitro and may well be members of a much larger class of regulators with HMG boxes; ii) the random-coil proteins  
25 HMGI4 and 17 with an as yet unknown function; iii) the HMGI-type proteins, which bind to the minor groove and include HMGI-C, HMGI, and HMGI-Y; the latter two are encoded by the same gene. It is of interest to note that published mapping positions of members of the HMG family coincide with  
30 published chromosome breakpoints of benign solid tumors such as those studied here. The HMGI(Y) gene, for instance, has been mapped to human chromosome 6p21 [69], which is known to be involved in recurrent translocations observed in uterine leiomyoma, lipoma, and pleomorphic salivary gland adenoma  
35 [76]. As listed in the Human Genome Data Base, not all known members of the HMG family have been chromosomally assigned yet, although for some of them a relatively precise mapping position has been established. For instance, HMGI7 to

chromosome 1p36.1-p35, HMG1L to 13q12, and HMG14 to 21q22.3; all chromosome segments in which chromosome breakpoints of the tumor types studied here have been reported [76].

Whether HMG1(Y) or any other of these HMG members are indeed  
5 affected in other subgroups of these tumors remains to be established. Of interest to mention, furthermore, are syndromes such as Bannayan-Zonana (McKusick #153480), Proteus (McKusick #176920), and Cowden (McKusick #158350); the latter syndrome is also called multiple hamartoma  
10 syndrome. In 60% of the individuals with congenital Bannayan-Zonana syndrome, a familial macrocephaly with mesodermal hamartomas, discrete lipomas and hemangiomas were found [70].

Finally, one aspect of our results should not  
15 escape attention. All the tumors that were evaluated in this study were of mesenchymal origin or contained mesenchymal components. It would be of great interest to find out whether the observed involvement of HMG1-C is mesenchyme-specific or may be found also in tumors of non-mesenchymal  
20 origin. The various DNA clones we describe here are valuable resources to address this important issue and should facilitate studies to conclusively implicate the HMG1-C gene in tumorigenesis.

## 25 EXAMPLE 3

Rearrangements of another member of the HMG gene family

### 1. Introduction

This example clearly demonstrates that within a  
30 given tumor entity (e.g. pulmonary chondroid hamartomas, uterine leiomyomas, endometrial polyps) tumors, histologically practically indistinguishable from each other, arise if either the HMG1-C gene or the HMG1(Y) gene is affected by chromosomal rearrangements. Thus, indeed a  
35 group of genes leading to aberrant mesenchymal growth including but not restricted to HMG1-C and HMG1(Y) can be defined.

## 2. Material and methods

### 2.1. Chromosome preparation

Chromosome preparation followed routine methods. Cells were treated with 30  $\mu$ l colcemide (10 $\mu$ g/ml) for 2-3 h and then harvested using the trypsin method (0.05% trypsin, 0.02% EDTA) followed by a hypotonic shock in six fold diluted medium TC 199 for 20 minutes at room temperature and methanol:acetic acid (3:1) fixation. Chromosomes were then GTG-banded.

10

### 2.2. In situ hybridization

In situ hybridisation was performed as outlined in one of the previous examples.

### 15 2.3. PAC library screening

The PAC library (Genome Systems Library Screening Service, St. Louis, Missouri, USA) was screened by PCR with a primer set specific for the HMGI(Y) gene. For screening we designed the forward primer with the sequence:

20 5'-CTC CAA GAC AGG CCT CTG ATG T-3' (intron 3)

and the reverse primer:

5'-ACC ACA GGT CCC CTT CAA ACT A-3' (intron 3)

giving rise to a fragment of 338 bp. For amplification the following thermal cycling was used: 94°C, 5 min, (94°C, 1 min, 59°C, 1 min, 72°C, 2 min) x 30, 72°C, 10 min.

25

### 2.4. DNA preparations from PAC clones

Bacterial colonies containing single PAC clones were inoculated into LB medium and grown overnight at 37°C. 660  $\mu$ l of the overnight culture were diluted into 25 ml of LB medium and grown to an OD<sub>550</sub> of 0.05-0.1. By addition of IPTG to a final concentration of 0.5 mM the P1 lytic replicon was induced. After addition of IPTG, growth was continued to an OD<sub>550</sub> of 0.5-1.5, and plasmid DNA was extracted using the alkaline lysis procedure recommended by Genome Systems.

30

35

### 3. Results

The primer set for screening the human PAC library was designed from sequences belonging to intron 3 of HMGI(Y). Because of sequence homology between HMGI-C and HMGI(Y) the amplified sequence of 338 bp was tested by homology search to be specific exclusively for HMGI(Y). Library screening resulted in three positive PAC clones that had an average insert length of approximately 100 kb. Two of these clones (Pac604, Pac605) were used for the following FISH studies. In order to prove if HMGI(Y) is rearranged in tumors with translocations involving 6p21.3 in either simple or complex form we performed FISH analysis on metaphase spreads from four primary pulmonary chondroid hamartomas and two endometrial polyps all with 6p21.3 abnormalities. For each case 20 metaphases were scored. At least one of the two PAC clones Pac604 and Pac605 described above was across the breakpoint in all six cases analyzed. These results clearly show that the breakpoints of the tumors with 6p21 aberrations investigated herein are clustered either within the HMGI(Y) gene or its close vicinity.

#### **EXAMPLE 4**

Hybrid HMGI-C in lipoma cells.

CDNA clones of the chromosome 3-derived lipoma-preferred partner gene LPP (>50 kb) were isolated and the nucleotide sequence thereof established. Data of a composite cDNA are shown in Fig. 4. An open reading frame for a protein (612 amino acids (aa)) with amino acid sequence similarity (over 50%) to zyxin of chicken was identified. Zyxin is a member of the LIM protein family, whose members all possess so-called LIM domains [78]. LIM domains are cysteine-rich, zinc-binding protein sequences that are found in a growing number of proteins with diverse functions, including transcription regulators, proto-oncogene products, and adhesion plaque constituents. Many of the LIM family members have been postulated to play a role in cell signalling and control of cell fate during development. Recently, it was demonstrated that LIM domains are modular

protein-binding interfaces [79]. Like zyxin, which is present at sites of cell adhesion to the extracellular matrix and to other cells, the deduced LPP-encoded protein (Fig. 6) possesses three LIM domains and lacks classical DNA-binding homeodomains.

In 3'-RACE analysis of Li-501/SV40, a HMGI-C containing fusion transcript was identified from which a hybrid protein (324 aa) could be predicted and which was subsequently predicted to consist of the three DBDs (83 aa) of HMGI-C and, carboxy-terminally of these, the three LIM domains (241 aa) encoded by LPP. In PCR analysis using appropriate nested amplicon sets similar HMGI-C/LPP hybrid transcripts were detected in various primary lipomas and lipoma cell lines carrying a t(3;12) and also in a cytogenetically normal lipoma. These data reveal that the cytogenetically detectable and also the hidden t(3;12) translocations in lipomas seem to result consistently in the in-phase fusion of the DNA-binding molecules of HMGI-C to the presumptive modular protein-binding interfaces of the LPP-encoded protein, thereby replacing the acidic domain of HMGI-C by LIM domains. Consequently, these protein-binding interfaces are most likely presented in the nuclear environment of these lipoma cells, where they might affect gene expression, possibly leading to aberrant growth control. Out of the large variety of benign mesenchymal tumors with chromosome 12q13-q15 aberrations, this is the first example of a chromosome translocation partner contributing recurrently and consistently to the formation of a well-defined tumor-associated HMGI-C fusion protein.

Figure 5 shows the cDNA sequence of the complete isolated LPP gene.

#### EXAMPLE 5

##### Diagnostic test for lipoma

A biopsy of a patient having a lipoma was taken. From the material thus obtained total RNA was extracted using the standard TRIZOL™ LS protocol from GIBCO/BRL as described in the manual of the manufacturer. This total RNA

was used to prepare the first strand of cDNA using reverse transcriptase (GIBCO/BRL) and an oligo dT(17) primer containing an attached short additional nucleotide stretch. The sequence of the primer used is as described in Example 2, under point 2.5.. RNase H was subsequently used to remove the RNA from the synthesized DNA/RNA hybrid molecule. PCR was performed using a gene specific primer (Example 2, point 2.5.) and a primer complementary to the attached short additional nucleotide stretch. The thus obtained PCR product was analysed by gel electrophoresis. Fusion constructs were detected by comparing them with the background bands of normal cells of the same individual.

In an additional experiment a second round of hemi-nested PCR was performed using one internal primer and the primer complementary to the short nucleotide stretch. The sensitivity of the test was thus significantly improved.

Figure 8 shows a typical gel.

#### EXAMPLE 6

Aberrations of 12q14-15 and 6p21 in pulmonary chondroid hamartomas

##### 1. Introduction

Pulmonary chondroid hamartomas (PCH) are often detected during X-ray examination of the lung as so-called coin lesions. However, lung metastases of malignant tumors and rarely lung cancers can also present as coin lesions. This example shows that FISH requiring a minimal amount of tumor cells can be used to correctly distinguish between the majority of PCHs and malignant tumors. Thus the test can successfully be applied e.g. to tumor cells obtained by fine needle aspiration.

##### 2. Materials and methods

Samples from a total of 80 histologically characterized PCHs were included in this study. Cell cultures, chromosome preparations and FISH were obtained or performed as described in the previous examples.

### 3. Results

Cytogenetic studies revealed that of the 80 PCHs studied cytogenetically 51 revealed detectable aberrations involving either 12q14-15 or 6p21. By FISH using either a pool of cosmids belonging to the HMGI-C gene or using the PAC clones of HMGI(Y) described in the previous example we were able to detect hidden structural rearrangements of those regions in 4 additional cases (3 with 12q and one with 6p involvement). Therefore, the FISH test alone can be used for a kit to precisely detect the rearrangement of either the HMGI-C or the HMGI(Y) gene rearrangements in more than 50% of the PCHs and is thus a valuable additional tool for the diagnosis of these tumors (without being restricted to this type of tumors as shown in two of the other examples).

15

#### **EXAMPLE 7**

Diagnosis of soft tissue tumors particularly of adipocytic origin

#### 1. Introduction

Adipocyte tissue tumors often cause diagnostic difficulties particularly when material taken from fine needle aspiration biopsies or cryosections has to be evaluated. This examples demonstrates the validity of the FISH test for the differential diagnosis of adipocyte tissue tumors and rare soft tissue tumors.

25

#### 2. Materials and methods

##### 2.1. Tumor samples

Tumor samples from three soft tissue tumors were investigated by FISH. Sample one (1) was from a adipocytic tumor and histologically it was either an atypical lipoma or a well-differentiated liposarcoma. The second case (tumor 2) was diagnosed to be most likely a myxoid liposarcoma but other types of malignant soft tissue tumors including aggressive angiomyxoma were also considered. The third tumor (tumor 3) was also of adipocytic origin and both a lipoma and a well differentiated liposarcoma were considered.

30  
35

## 2.2. Isolation of cells and FISH

The tumor samples were enzymatically disaggregated following routine methods. The resulting single cell suspensions were centrifuged and the suspensions were fixed using methanol:glacial acetic acid (3:1) at room temperature for 1 hour. The cell suspensions were then dropped on clean dry slides and allowed to age for 6 hours at 60°C. FISH was performed using molecular probes from the HMGI-C gene as described in the previous examples.

10

## 3. Results

At the interphase level tumor 1 and 2 both showed split signals for one of the alleles. These findings are compatible with the diagnosis of benign tumors i.e. an atypical lipoma in the first case and an aggressive angiomyxoma in the second case. They allowed to rule out the presence of malignant adipocytic tissue tumors.

In the third case the FISH revealed a high degree of amplification of the MAR region or part of it. Since the amplification units observed in giant marker or ring chromosomes in well-differentiated liposarcomas can involve the MAR region these findings leads to the diagnosis of a well-differentiated liposarcoma. The three cases presented within this example show the usefulness of the DNA probes described. They can be used in a kit for a relatively simple and fast interphase FISH experiment offering an additional tool for the diagnosis of soft tissue tumors.

## EXAMPLE 8

Expression of the HMGI-C gene in normal tissue

### 1. Introduction

It is the aim of this example to show that the expression of the HMGI-C gene is mainly restricted to human tissues during embryonic and fetal development. In contrast, in most normal tissues of the adult particularly, including those tissues and organs tumors with HMGI-C rearrangements can arise from, no expression can be noted. This indicates that even the transcriptional re-activation of the gene can

initiate tumorigenesis. On the other hand it underlines the usefulness of antisense strategies (including those antisense molecules directed towards the normal HMGI-C mRNA) to inhibit or stop tumor growth.

5

## 2. Materials and methods

### 2.1. Tissue samples

All adult tissue samples used for this study were taken from surgically removed tissue frozen in liquid nitrogen within a period of 15 min after removal. Most of the samples were from adjacent normal tissue removed during tumor surgery. In detail we have used 8 samples taken from fat tissues at various anatomical sites, 20 samples taken from myometrial tissue, 8 samples taken from lung tissue, 4 samples taken from the salivary glands (Glandula parotis and Glandula submandibularis), one tissue sample taken from the heart muscle, 25 samples taken from breast tissue from patients of different ages, 2 samples from the brain, 3 liver samples, 7 samples taken from renal tissue, and 20 embryonic/fetal tissue (extremities, 6 samples) from embryos/fetuses (10-14th gestational week) after abortion from socio-economic reasons.

In addition, three cell lines were used: As a control for HMGI-C expression we used the hepatoma cell line Hep 3B and the cell line L14 established from a lipoma with the typical translocation t(3;12). HeLa cells were used as a negative control because RT experiments reproduced for 10 times did not reveal HMGI-C expression in our own studies.

### 2.2. RT-PCR for the expression of HMGI-C

100 mg of tissue sample was homogenized, and RNA was isolated using the trizol reagent (GibcoBRL, Eggenstein, Germany) containing phenol and isothiocyanate. cDNA was synthesized using a poly(A)-oligo(dt)17 primer and M-MLV reverse transcriptase (GibcoBRL, Eggenstein, Germany). Then, a hemi-nested PCR was performed.

For first and second PCR the same lower primer (Revex 4) (5'-TCC TCC TGA GCA GGC TTC-3' (exon 4/5)) was

used. In the first round of PCR the specific upper primer (SE1) (5'-CTT CAG CCC AGG GAC AAC-3' (exon 1)), and in the second round of PCR the nested upper primer (P1) (5'-CGC CTC AGA AGA GAG GAC-3' (exon 1)) was used. Both rounds of PCR were performed in a 100  $\mu$ l volume containing 10 mM Tris/HCl pH 8.0, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.001 % gelatin, 100  $\mu$ M dATP, 100  $\mu$ M dTTP, 100  $\mu$ M dGTP, 100  $\mu$ M dCTP, 200 nM upper primer, 200 nM lower primer, and 1 unit/100  $\mu$ l AmpliTaq polymerase (Perkin Elmer, Weiterstadt, Germany).

10 Amplification was performed for 30 cycles (1 min 94°C, 1 min 53°C, 2 min 72°C). As template in the first round of PCR cDNA derived from 250 ng total RNA, and in the second round of PCR 1  $\mu$ l of the first PCR reaction mix was used.

### 15 2.3. Control assay for intact mRNA/cDNA

As control reaction for intact RNA and cDNA PCR a test based on the amplification of the cDNA of the housekeeping gene glyceraldehyde 3-phosphate dehydrogenase (GAPDH). PCR reaction was performed for 35 cycles under the same conditions as described above for the first round of PCR of HMGI-C expression.

20

### 3. Results

As for the expression studies all experiments were repeated at least twice. To assure that all RNA and cDNA preparations used for the RT-PCRs were intact (otherwise resulting in false negative results) routinely a RT-PCR for expression of the housekeeping gene GAPDH was performed. A positive GAPDH RT-PCR results in a 299 bp fragment. Only samples revealing a positive GAPDH RT-PCR were included in this study. As the result of RT-PCR in HMGI-C positive cells such as Hep 3B and L14 a specific 220 bp fragment is detectable. HeLa cells did not show an expression of HMGI-C. Except for two myometrial samples (most likely due to myomas at a submicroscopic level) all normal tissue samples taken from adult individuals did not show any detectable level of HMGI-C expression. In contrast, all fetal/embryonic tissues tested revealed HMGI-C expression.

25

30

35

**EXAMPLE 9**

Expression of the HMGI-C gene as a diagnostic tool for the early detection of leukemias

**1. Introduction**

5 Cytogenetically detectable aberrations affecting the HMGI-C gene have been found in a variety of benign solid tumors of mesenchymal origin. Apparently, the aberrations also lead to the transcriptional activation of the gene. Since blood cells are also of mesenchymal origin, it was  
10 tempting to check leukemic cells for HMGI-C expression. The present example shows that the activation of the gene in cells of the peripheral blood is a suitable marker indicating immature cells/abnormal stem cells found in leukemias. Since the expression of HMGI-C can be determined  
15 with a high degree of sensitivity the RT-PCR for the expression of the gene can be used for a very early detection of various hematological diseases.

**2. Materials and methods**

20 Samples from peripheral blood of 27 patients with different types of leukemias including 19 patients with Philadelphia-chromosome positive CML, 5 patients with AML, and 3 patients with ALL were used for determination of HMGI-C expression. Blood samples from 15 healthy probands served  
25 as controls.

RT-PCR for the expression of HMGI-C was performed as outlined in example 8.

**3. Results**

30 Whereas expression of HMGI-C was clearly detectable in all blood samples from leukemic patients there was no expression noted in any of the blood samples taken from the control persons. There is no evidence that the transcriptional activation of the gene is due to mutations  
35 affecting the gene or its surroundings. It is more reasonable to assume that the activation is rather a secondary effect related to the immaturity of the cells or their abnormal proliferation. However, the high and even

improvable sensitivity makes e.g. a kit based on the RT-PCR for the expression of the HMGI-C gene a very suitable diagnostic tool.

## 5 EXAMPLE 10

The transcriptional re-expression of the HMGI-C gene can lead to the initiation of the tumors

### 1. Introduction

This example clearly shows that for some tumor entities chromosomal breakpoints located 5' of the HMGI-C gene do also exist indicating that the transcriptional up-regulation of the gene is sufficient to initiate growth of the corresponding tumor types.

## 15 2. Materials and methods

### 2.1. Cell Culture

After surgery the tumor samples (three pulmonary chondroid hamartomas, one uterine leiomyoma) were washed with Hank's solution supplemented with penicillin (200 IU/ml) and streptomycin (200 µg/ml). Tumors were disaggregated with collagenase for 5-6 h at 37°C. The suspension containing small fragments and single cells was resuspended in culture medium TC 199 with Earle's salts supplemented with 20% fetal bovine serum, 200 IU/ml penicillin, and 200 µg/ml streptomycin.

### 2.2. Chromosome preparations

Chromosome preparation followed routine methods. Cells were treated with 30 µl colcemide (10 µg/ml) for 2-3 h and then harvested using the trypsin method (0.05% trypsin, 0.02% EDTA) followed by a hypotonic shock in six fold diluted medium TC 199 for 20 minutes at room temperature and methanol:acetic acid (3:1) fixation. Chromosomes were then GTG-banded.

35

### 2.3. FISH studies

To identify the chromosomes unambiguously, FISH was performed after GTG-banding of the same metaphase

spreads. As DNA probes we used five cosmids belonging to a YAC-contig overspanning the HMGI-C gene as described in the Legenda of figure 2. Three of these cosmids (27E12, 185H2, 142H1) are mapping to the third intron of HMGI-C, whereas 5 cosmids 260C7 and 245E8 are localized at the 3' or the 5' end respectively. The slides were analyzed using a Zeiss (Zeiss, Oberkochen, Germany) Axioplan fluorescence microscope. Results were processed and recorded with the Power Gene Karyotyping System (PSI, Halladale, Great Britain). Rapid amplification of cDNA ends (RACE) was performed as described in one of the former examples.

### 3. Results

All four tumors showed the same type of cytogenetic abnormality, i.e. the presence of 47 chromosomes including two apparently normal chromosomes 12 and an additional derivative 14 der(14)t(12;14)(q14-15;q24) but without a corresponding der(12). Since the 3'-5' orientation of the HMGI-C is towards the centromere a single break within the HMGI-C gene would have led to the loss of its 5' part along with the loss of the der(12). We have therefore performed a series of FISH experiments in order to determine the breakpoints more precisely. Using the five cosmids 260C7, 27E12, 185H2, 142H1, and 245E8 hybridization signals of the same intensity were observed at both normal chromosomes 12 and at the additional der(14). The FISH results revealed that in all four cases chromosomal breakpoints were located 5' of the HMGI-C gene.

The breakpoint assignment in all four cases 5' of the HMGI-C gene fits well with the results of the RACE-PCR. In addition to the normal HMGI-C transcripts we were able to detect aberrant transcripts in all three tumors. Sequences showed that they were not derived from chromosome 14 but from intron 3 of HMGI-C probably due to cryptic splice sites. However, the RACE results revealed that there was indeed HMGI-C expression in all four cases.

**EXAMPLE 11****Re-differentiation of leukemic cells****1. Introduction**

Expression of the HMGI-C gene is frequently  
5 strongly elevated in a wide variety of tumors, solid tumors  
as well as leukemias. It was speculated that the HMGI-C  
protein might play a key role in transformation of cells.  
This example shows that expression of the HMGI-C gene can be  
strongly reduced by expressing antisense HMGI-C sequences  
10 and that reduction of HMGI-C levels in tumor cells results  
in reversion of the transformed phenotype. Thus the  
expression or administration of antisense molecules can be  
successfully applied therapeutically.

**2. Materials and methods****2.1. Tumor cell lines**

Tumor cell lines were generated from a primary  
malignant salivary gland tumor and a primary breast  
carcinoma. Cell lines were established as described by  
20 Kazmierczak, B., Thode, B., Bartnitzke, S., Bullerdiek, J.  
and Schloot, W., "Pleomorphic adenoma cells vary in their  
susceptibility to SV40 transformation depending on the  
initial karyotype.", Genes Chrom. Cancer 5:35-39 (1992).

**2.2. Assay of the transformed state**

Soft agar colony assays were performed as  
described by Macpherson and Montagnier, "Agar suspension  
culture for the selective assays of cells transformed by  
polyoma virus." Virology 23, 291-294 (1964).

30 Salivary gland and breast tumor cells were  
propagated in TC199 culture medium with Earle's salts,  
supplemented with 20% fetal bovine serum (GIBCO), 200 IU/ml  
penicillin, and 200 µg/ml streptomycin.

Tumorigenicity of the transfected salivary gland  
35 (AD64) and breast cell lines was tested by injecting cells  
subcutaneously into athymic mice.

### 2.3. Transfection assay

Transfections were performed using various protocols, namely:

1. The calcium phosphate procedure of Graham and Van der Eb ("A new technique for the assay of the infectivity of human adenovirus." Virology 52, 456-467 (1973) ).
2. Lipofection: Transfections were carried out using liposome-mediated DNA transfer (lipofectamine, GibcoBRL) according to the guidelines of the manufacturer.

10

### 2.4. Antisense constructs

- Sense and antisense constructs of the HMGI-C gene were obtained by inserting human HMGI-C cDNA sequences in both the sense and antisense orientation in expression vectors under the transcriptional control of various promoter contexts, e.g. the long terminal repeat of Moloney murine leukemia virus, a CMV promoter, or the early promoter of SV40. For example, the CMV/HMGI-C plasmid was constructed by cloning a human HMGI-C cDNA fragment containing all coding sequences of human HMGI-C in pRC/CMV (Invitogen) allowing expression under control of the human cytomegalovirus early promoter and enhancer, and selection for G418 resistance.

## 25 3. Results

### 3.1. Reversion of the transformed phenotype

- Reversion of the transformed phenotype was observed in breast and salivary gland tumors cells after induction of antisense HMGI-C expression in these tumor cells. A strong reduction in tumorigenicity was observed as measured by soft agar colony assay and in vivo in athymic mice. Immunoprecipitation and Western blot analysis indicated a strong reduction of HMGI-C protein levels in the cells expressing antisense HMGI-C sequences. Therefore, this approach can be used therapeutically in tumors with involvement of HMGI-C.

**EXAMPLE 12**

Animal tumor models involving HMGI-C as tools in in vivo therapeutic drug testing.

On the basis of the acquired HMGI-C knowledge,  
5 animal tumor models can be developed as tools for in vivo drug testing. To achieve this objective (for instance for uterine leiomyoma), two approaches can be used, namely gene transfer (generation of transgenic animals) on the one hand and gene targeting technology (mimicking in vivo of a  
10 specific genetic aberration via homologous recombination in embryonic stem cells (ES cells)) on the other.

These technologies allow manipulation of the genetic constitution of complex living systems in specific and pre-designed ways. For extensive technical details, see  
15 B. Hogan, R. Beddington, F. Constantini, and E. Lacy; In: Manipulating the mouse embryo, A Laboratory Manual. Cold Spring Harbor Press, 1994; ISBN 0-87969-384-3.

To aim at the inactivation or mutation of the HMGI-C gene, specifically in selected cell types and  
20 selected moments in time, the recently described Cre/LoxP system can be used (Gu, H. et al. Deletion of a DNA polymerase  $\beta$  gene segment in T cells using cell type-specific gene targeting. Science 265, 103-106, 1994). The Cre enzyme is a recombinase from bacteriophage P1 whose  
25 physiological role is to separate phage genomes that become joined to one another during infection. To achieve so, Cre lines up short sequences of phage DNA, called loxP sites and removes the DNA between them, leaving one loxP site behind. This system has now been shown to be effective in mammalian  
30 cells in excising at high efficiency chromosomal DNA. Tissue-specific inactivation or mutation of a gene using this system can be obtained via tissue-specific expression of the Cre enzyme.

As an example, the development of animal model  
35 systems for uterine leiomyoma using a member of the MAG gene family will be outlined below, such that the models will be instrumental in in vivo testing of therapeutic drugs.

Two approaches may be followed:

- a) in vivo induction of specific genetic aberrations as observed in human patients ((conditional) gene (isogenic) targeting approach); and
- b) introduction of DNA constructs representative for the genetic aberrations observed in patients (gene transfer approach).

DNA constructs to be used in gene transfer may be generated on the basis of observations made in patients suffering from uterine leiomyoma as far as structure and expression control are concerned; e.g. HMGI-C fusion genes with various translocation partner genes, especially the preferential translocation partner gene of chromosome 14 located in the YAC contig represented by CEPH YACs 6C3, 89C5, 308H7, 336H12, 460A6, 489F4, 902F10, 952F5, 958C2, 961E1, and 971F5, truncated genes encoding basically the three DNA binding domains of HMGI-C, and complete HMGI-C or derivatives of HMGI-C under control of a strong promoter.

#### EXAMPLE 13

- 20 The preparation of antibodies against HMGI-C

One type of suitable molecules for use in diagnosis and therapy are antibodies directed against the MAG genes. For the preparation of rabbit polyclonal antibodies against HMGI-C use was made of the following

- 25 three commercially available peptides:

(H-ARGEGAGQPSTSAQGQPAAPAPQKR) 8-Multiple Antigen Peptide

(H-SPSKAAQKKAEATGEKR) 8-MAP

(H-PRKWPQQVVQKKPAQEE) 8-MAP

obtainable from Research Genetics Inc., Huntsville, AL, USA.

- 30 The polyclonal antibodies were made according to standard techniques.

## TABLES

TABLE I  
ANALYSIS OF YAC CLONES

CEPH-Code	Size (kb)	Landmark left	#	Landmark right	#	Chimeric
183F3	715			[RM10]		YES (L + R)
70E1	450	RM29	U27125			ND
95F1	390			RM30	U29054	ND
201H7	320	RM13	U29051	RM14	U29053	ND
186G12	320					ND
354B6	280					YES (R)
126G8	410					ND
258F11	415	RM4	U29052			ND
320F6	290	RM5	U29050	RM21	U29047	ND
234G11	475	RM7	U29046			ND
375H5	290					ND
252E10	510	[RM15]		RM16	U29048	YES (L)
181C8	470			RM26	U29045	ND
107D1	345	RM31	U29043			ND
499C5	320	RM44	U29044	RM46	U29037	ND
340B6	235					ND
532C12	400	RM45	U29041			ND
138C5	510	[RM59]		RM65	U29042	YES (L)
145F2	490	RM60	U29030	RM66	U29040	ND
106E8	340	RM57	U29033	RM63	U29038	ND
55G1	365	RM56	U29031	RM62	U29039	ND
103G7	370	RM85	U29025	RM80	U29036	ND
295B10	295	RM77	U29035	RM81	U29026	ND
338C2	200	RM78	U29034	RM82	U29029	ND
391C12	160	[RM79]		RM83	U29027	YES (L)
476A11	225	[RM87]		RM84	U29032	YES (L)
138F3	460	RM90	U29028	RM91	U29019	ND
226E7	500	RM48	U29024	RM54	U29015	ND
499E9	375	RM51	U29016			YES (R)
312F10	580	[RM50]		RM69	U29021	YES (L)
825G7	950					ND
34B5	315	RM88	U29020	RM89	U29013	ND
94A7	610					YES (R)
305B2	660					YES (L)
379H1	280	RM104	U29014	RM105	U29009	ND
444E6	350	RM92	U29017	RM93	U29010	ND
446H3	370	RM94	U29011	RM95	U29018	ND
403B12	380					ND
261E5	500	RM102	U29012	RM103	U26689	ND
78B11	425					ND
921B9	1670					ND
939H2	1750					ND
188H7	360					ND
142F4	390					ND
404E12	350					ND
164A3	375					ND
244B12	415	RM106	U29007	RM107	U29008	ND
275H4	345	RM108	U29004	RM109	U29005	ND
320F9	370					ND
51F8	450					ND
242A2	160	CH1	U29006			ND

TABLE I (continued)  
ANALYSIS OF YAC CLONES

253H1	400					ND
303F11	320					ND
322C8	410			CH2	U29002	ND
208G12	370	RM96	U29002	RM97	U27135	ND
341C1	270	RM98	U26647	RM99	U27130	ND
354F1	270					ND
452E1	270	CH5	U27136			ND
41A2	310					ND
934D2	1370					ND
944E8	1290			CH8	U25792	ND
2G11	350					ND
755D7	1390					YES (L)
365A12	370					ND
803C2	1080					ND
210C1	395	RM70	U28998	RM86	U27133	ND
433C8	360	RM73	U29000	RM76	U27132	ND
402A7	500	RM41	U28994	[RM42]		YES (R)
227E8	465	RM53	U27134	RM55	U22996	ND
329F9	275	RM72	U28793	RM75	U22997	ND
261E6	395	[RM71]		RM74	U22995	YES (L)
348F2	370			[RM136]		YES (R)
6F3	320	RM35	U27140	RM36	U27141	
59F12	430	RM34	U28794	RM33	U27131	
265H3	300			RM40	U26999	

YAC clones were isolated from CEPH YAC libraries as described in Materials and Methods. ND: not detected by methods used. Landmarks not mapping within the 6 Mb contig have been bracketed. GenBank accession numbers are given (#).

TABLE II  
PCR Primers

STS name (STS 12-)	Nucleotide sequence 5'-3'	Product size (bp)	T <sub>m</sub> (°C)
CH1	TGGGACTTAACGGATTTTCGA TGTGGTTTCATTTCATGCATTA	213	58
CH2	TCCATCATCATCTCAAAACA CTCTACCAATATGGATATAACAG	145	58
CH5	GCAGCTCAGGCTCCTTCCCA TGGCTTCTGTGAAACGCGAGA	143	58
CH8	TCTCCACTGCTTCCATTTCAC ACACAAAACCACTGGGGGCT	147	58
CH9	CAGCTTTGGAATCAGTGAGG CCTGGGGAGAGAGGAGTAAAG	262	58
RM1	GAGCTTCCTATCTCATCC ATGCTTGTGTGTGAGTGG	308	60
RM4	TTTGCTAAGCTAGGTGCC AGCTTCLAGACCCATGAG	235	60
RM5	CAGTTCTGAGACTGCTTG TAATAGCAGGGACTCAGC	324	60
RM7	CTTGTCTCAATTCCTTTAAGGG CACCCTTTTATAGATCCTAC	538	58
RM13	GAATGTTTCATCACAGTGCTG ATGTGAGGTTCTGCTGAG	±500	58
RM14	TTCTCATGGGGTAAGGACAG AAAGCTGCTTATATAGGGAATC	158	58
RM16	CCTTGGCTTAGATATGATACAC GCTCTTCAGAAATATCCTATGG	252	58
RM21	CCTTAGCAGTTGCTTGTCTG TCGTCACAGGACATAGTCAC	290	58
RM26	TCTATGGTATGTTATACAGATG CAGTGAGATCCTGTCTCTA	102	58
RM31	TCTGTGATGTTTTAAGCCACTTAG AATCTGTGTCCCTGCCACC	239	56
RM33	AATCTTCCTCACCTCCCACC AATCTGCAGAGAGGTCCAGC	±600	60
RM34	AATCTCCATCTGGGCCTGG GAACGCTAAGCATGTGGGAG	±600	60
RM36	CTCCAACCATGGTCCAAAC GACCTCCAGTGGCTCTTTAG	296	60
RM46	ACCATCAGATCTGGCACTGA TTACATTGGAGCTGTCTATGC	241	57
RM48	TCCAGGACATCCTGAAAATG AGTATCCTGCCACTTCTGCCAG	391	58
RM51	GATGAATCTGAGGTGCCCTTC TCAPACCCAGCTTTGACTCC	311	60
RM53	GTCTTCAAAAACGCTTTCTCTG TGGTTTGCATAATGGTGATG	333	60

TABLE II (continued)

## PCR Primers

STS name (STS 12-)	Nucleotide sequence 5'-3'	Product size (bp)	T <sub>m</sub> (°C)
RM60	TACACTACTCTGCAGCAGAC	94	58
RM69	TCTGAGTCAATCACATGTCC CTCCCCAGATGATCTCTTTC CGGTAGGAAATAAAGGAGAG	236	58
RM72	TATTTACTAGCTGGCCCTGG CATCTCAGGCACACACAATG	101	62
RM76	ATTGAGAGAGTGGCCAACT GGGATAGGTCTTCTGCAATC	496	58
RM85	TCCAACATACTGAGTGACC TCCATTTCACTGTAGCACTG	435	58
RM86	GTAATCAACCATTCCCCTGA AAATAGCTGGTATGGTGGC	203	56
RM90	ACTGCTCTAGTTTTCAAGGA AATTTACCTGACAGTTTCTC	257	58
RM93	GCATTTGACGTCCAATATTG ATTCCATTGGCTAACACAAG	347	60
RM98	GCAAACTTTGACTGAAACG CACAGAGTATCGCACTGCAT	356	58
RM99	AAGAGATTTCCCATGTTGTG CTAGTGCCTTCACAAGAACC	240	58
RM103	AATTCCTTGAGGGGTTCACTG TCCACACTGAGAGCTTTTTCA	199	60
RM108	GTGGTTCTGTACAGCAGTGG TGAGAAAATGCTGCCAAT	439	60
RM110	GCTCTACCAGGCATACAGTG ATTCTAGCATCTTTTCACG	328	58
RM111	ATATGCATTAGGCTCAACCC ATCCCAAGGTCAACATGAC	312	58
RM130	ATCCTTACATTTCCAGTGGCATTCA CCCAGAAGACCCACATTCCTCAT	336	58
RM131	TTTTAAGTTTCTCCAGGGAGGAGAC AATAGGCTCTTTGGAAAGCTGGAGT	226	58
RM132	TCTCAGCTTAATCCAAGAAGGACTTC GGCATATTCTCAACAATTTATGCTT	376	58
RM133	TGGAGAAGCTATGGTGCTTCCTATG TGACAAATAGGTGAGGGAAGTTGTTAT	225	58
EST01096	TCACACGCTGAATCAATCTT CAGCAGCTGATACAAGCTTT	186	58
IFNG	TGTTTTCTTTCCGATAGGT CTGGGATGCTCTTCGACCTC	150	52
Rap1B	CCATCCAACATCTTAATGGAC CAGCTGCAAACTCTAGGACTATT	149	58

STSs were isolated as described in Materials and Methods, or retrieved from literature for EST01096, IFNG, and Rap1B.

Table 3

Genome Data Base accession numbers (D-numbers) of the various sequences indicated in Figure 1.

## Genome Data Base

per	locus_symbol	per	per_gdb_id	(41 rows affected)	locus	locus_gdb_id
CH1-lower/CH1-upper	D12S1484	CH1-lower/CH1-upper	G00-595-292		D12S1484	G00-595-415
CH2-lower/CH2-upper	D12S1485	CH2-lower/CH2-upper	G00-595-295		D12S1485	G00-595-416
CH5-lower/CH5-upper	D12S1486	CH5-lower/CH5-upper	G00-595-298		D12S1486	G00-595-417
CH8-lower/CH8-upper	D12S1487	CH8-lower/CH8-upper	G00-595-301		D12S1487	G00-595-418
CH9-lower/CH9-upper	D12S1488	CH9-lower/CH9-upper	G00-595-304		D12S1488	G00-595-419
EH2-lower/CH2-upper	D12S1489	EH2-lower/CH2-upper	G00-595-307		D12S1489	G00-595-420
EH3-lower/CH3-upper	D12S1490	EH3-lower/CH3-upper	G00-595-310		D12S1490	G00-595-421
EH4-lower/CH4-upper	D12S1491	EH4-lower/CH4-upper	G00-595-313		D12S1491	G00-595-422
RH13-lower/RH13-upper	D12S1492	RH13-lower/RH13-upper	G00-595-316		D12S1492	G00-595-423
RH14-lower/RH14-upper	D12S1493	RH14-lower/RH14-upper	G00-595-319		D12S1493	G00-595-424
RH16-lower/RH16-upper	D12S1494	RH16-lower/RH16-upper	G00-595-322		D12S1494	G00-595-425
RH25-lower/RH25-upper	D12S1495	RH25-lower/RH25-upper	G00-595-325		D12S1495	G00-595-426
RH26-lower/RH26-upper	D12S1496	RH26-lower/RH26-upper	G00-595-328		D12S1496	G00-595-427
RH31-lower/RH31-upper	D12S1497	RH31-lower/RH31-upper	G00-595-331		D12S1497	G00-595-428
RH33-lower/RH33-upper	D12S1498	RH33-lower/RH33-upper	G00-595-334		D12S1498	G00-595-429
RH34-lower/RH34-upper	D12S1499	RH34-lower/RH34-upper	G00-595-337		D12S1499	G00-595-430
RH36-lower/RH36-upper	D12S1500	RH36-lower/RH36-upper	G00-595-340		D12S1500	G00-595-431
RH46-lower/RH46-upper	D12S1501	RH46-lower/RH46-upper	G00-595-343		D12S1501	G00-595-432
RH48-lower/RH48-upper	D12S1502	RH48-lower/RH48-upper	G00-595-346		D12S1502	G00-595-433
RH51-lower/RH51-upper	D12S1503	RH51-lower/RH51-upper	G00-595-349		D12S1503	G00-595-434
RH53-lower/RH53-upper	D12S1504	RH53-lower/RH53-upper	G00-595-352		D12S1504	G00-595-435
RH60-lower/RH60-upper	D12S1505	RH60-lower/RH60-upper	G00-595-355		D12S1505	G00-595-436
RH69-lower/RH69-upper	D12S1506	RH69-lower/RH69-upper	G00-595-358		D12S1506	G00-595-437
RH72-lower/RH72-upper	D12S1507	RH72-lower/RH72-upper	G00-595-361		D12S1507	G00-595-438
RH76-lower/RH76-upper	D12S1508	RH76-lower/RH76-upper	G00-595-364		D12S1508	G00-595-439
RH85-lower/RH85-upper	D12S1509	RH85-lower/RH85-upper	G00-595-367		D12S1509	G00-595-440
RH86-lower/RH86-upper	D12S1510	RH86-lower/RH86-upper	G00-595-370		D12S1510	G00-595-441
RH90-lower/RH90-upper	D12S1511	RH90-lower/RH90-upper	G00-595-373		D12S1511	G00-595-442
RH93-lower/RH93-upper	D12S1512	RH93-lower/RH93-upper	G00-595-376		D12S1512	G00-595-443
RH98-lower/RH98-upper	D12S1513	RH98-lower/RH98-upper	G00-595-379		D12S1513	G00-595-444
RH99-lower/RH99-upper	D12S1514	RH99-lower/RH99-upper	G00-595-382		D12S1514	G00-595-445
RH29-lower/RH29-upper	D12S1496	RH29-lower/RH29-upper	G00-595-385		D12S1515	G00-595-446
RH103-lower/RH103-upper	D12S1516	RH103-lower/RH103-upper	G00-595-388		D12S1516	G00-595-447
RH108-lower/RH108-upper	D12S1517	RH108-lower/RH108-upper	G00-595-391		D12S1517	G00-595-448
RH110-lower/RH110-upper	D12S1518	RH110-lower/RH110-upper	G00-595-394		D12S1518	G00-595-449
RH111-lower/RH111-upper	D12S1519	RH111-lower/RH111-upper	G00-595-397		D12S1519	G00-595-450
RH121-lower/RH121-upper	D12S1520	RH121-lower/RH121-upper	G00-595-400		D12S1520	G00-595-451
RH130-lower/RH130-upper	D12S1521	RH130-lower/RH130-upper	G00-595-403		D12S1521	G00-595-452
RH131-lower/RH131-upper	D12S1522	RH131-lower/RH131-upper	G00-595-406		D12S1522	G00-595-453
RH132-lower/RH132-upper	D12S1523	RH132-lower/RH132-upper	G00-595-409		D12S1523	G00-595-454
RH133-lower/RH133-upper	D12S1524	RH133-lower/RH133-upper	G00-595-412		D12S1524	G00-595-455

TABLE 4

FISH mapping of chromosome 12 breakpoints in primary benign solid tumors  
to a subregion of MAR

Tumor type	Breakpoint within MAR	Fraction of tumors with breakpoints within main breakpoint cluster region*
Lipoma	6/6	6/6
Pleomorphic salivary gland adenoma	7/7	5/7
Uterine leiomyoma	7/8	7/8
Hamartoma of the breast	1/1	1/1
Fibroadenoma of the breast	1/1	1/1
Hamartoma of the lung	8/9	8/9
Angiomyxoma	1/1	1/1

\* Tumor samples were collected and analyzed at the histopathology and cytogenetics facilities of the University of Bremen. A mixture of cosmid clones 27E12 and 142H1 was used as molecular probe in FISH analysis.

Table 5-1

Clone	Tumor/Cell Line	Diversification Point	Nuc. Sequences (10b)	Chrom. sources RACE	poly(A) signal	#A's primer-set
pCH108	LI-14/SV40	exon/Intron 3 (DBD3)	TAGGAATGG GTGAGTAAATA	12 ?		
pCH113	LI-538/SV40	after DBD3	TAGGAATGG AATACTCTGG	12 ?	3q27	26 ?
pCH1234	#2528-90	after DBD3	TAGGAATGG AATACTCTGG	12 ?	3q27	26 ?
pCH1259	#2344-94	after DBD3	TAGGAATGG AATACTCTGG	12 ?		17
pCH1260	#2344-94	after DBD3	TAGGAATGG AATACTCTGG	12 ?		?
pCH148	#192	after DBD3	TAGGAATGG AATACTCTGG	12 ?		?
pCH145	#568-92	after DBD3	TAGGAATGG AATACTCTGG	12 N.T.		---
pCH1247	#568-92	after DBD3	TAGGAATGG AATACTCTGG	12	12	18
pCH1261	#2344-94	after DBD3	TAGGAATGG AATACTCTGG	12		17
pCH1212	#1321-89 ?	after DBD3	TAGGAATGG AATACTCTGG	12		53
pCH1213	#1321-89 ?	after DBD3	TAGGAATGG AATACTCTGG	12		27
pCH1228	#1321-89	after DBD3	TAGGAATGG AATACTCTGG	12		16
pCH216	#2100-89	after DBD3	TAGGAATGG AATACTCTGG	12		?
pCH1229	#2100-89	after DBD3	TAGGAATGG AATACTCTGG	12		13
pCH1232	#2100-89	after DBD3	TAGGAATGG AATACTCTGG	12		12
pCH177	#367	after DBD3	TAGGAATGG AATACTCTGG	12		25
pCH191	#25 ?	after DBD3	TAGGAATGG AATACTCTGG	12		29
pCH210	#837-88	after DBD3	TAGGAATGG AATACTCTGG	12		177
pCH1211	#837-88	after DBD3	TAGGAATGG AATACTCTGG	12		(177)
pCH230	#2100-89	after DBD3	TAGGAATGG AATACTCTGG	12		22
pCH1233	#2100-89	after DBD3	TAGGAATGG AATACTCTGG	12		22
pCH111	LI-501/SV40	after DBD3	TAGGAATGG AATACTCTGG	12		27
pCH112	LI-501/SV40	after DBD3	TAGGAATGG AATACTCTGG	12		18
pCH114	LI-538/SV40	after DBD3	TAGGAATGG AATACTCTGG	12		33
pCH115	LI-538/SV40	after DBD3	TAGGAATGG AATACTCTGG	12		33
pCH147	#192	after DBD3	TAGGAATGG AATACTCTGG	12		24
pCH153	#203	after DBD3	TAGGAATGG AATACTCTGG	12		115-AB
pCH169	#294	after DBD3	TAGGAATGG AATACTCTGG	12		147
pCH172	#294	after DBD3	TAGGAATGG AATACTCTGG	12		---
pCH173	#294	after DBD3	TAGGAATGG AATACTCTGG	12		169
pCH174	#294	after DBD3	TAGGAATGG AATACTCTGG	12		172-AB
pCH110	LN-30.1/SV40	after DBD3	TAGGAATGG AATACTCTGG	12		173-AB
pCH164	Myo168.1	after DBD3	TAGGAATGG AATACTCTGG	12		174-AB
pCH165	Myo168.1	after DBD3	TAGGAATGG AATACTCTGG	12		110 AB
pCH168	Myo196.4	after DBD3	TAGGAATGG AATACTCTGG	12		164-AB
pCH209	#837-88 ?	after DBD3	TAGGAATGG AATACTCTGG	12		165-AB
pCH217	#2100-89 ?	after DBD3	TAGGAATGG AATACTCTGG	12		168
pCH219	#2100-89 ?	after DBD3	TAGGAATGG AATACTCTGG	12		---
pCH220	#3391-90	after DBD3	TAGGAATGG AATACTCTGG	12		172-AB
pCH1223	#CG592	after DBD3	TAGGAATGG AATACTCTGG	12		173-AB
pCH226	#3100-88	after DBD3	TAGGAATGG AATACTCTGG	12		174-AB
pCH1231	#2100-89	after DBD3	TAGGAATGG AATACTCTGG	12		15
pCH1246	#568-92	after DBD3	TAGGAATGG AATACTCTGG	12		17
pCH1248	#568-92	after DBD3	TAGGAATGG AATACTCTGG	12		30
pCH1249	#2617-93	after DBD3	TAGGAATGG AATACTCTGG	12		16
pCH1251	#2617-93	after DBD3	TAGGAATGG AATACTCTGG	12		30
pCH1250	#2617-93	after DBD3	TAGGAATGG AATACTCTGG	12		23
pCH1252	#2617-93	after DBD3	TAGGAATGG AATACTCTGG	12		19
						13
						13

Table 5-2

Clone	Tumor/Cell Line	Diversion Point	Nuc. Sequences (10b)	Chrom. sources RACE cylogen.	poly(A) signal	#A's primer- set
exon/intron 4 (SPACER)						
PC1109	LM30.1/SV40	after spacer	GCCTGCTCAG	12	AATAAA	17
PC1138	#278-93	after spacer	GCCTGCTCAG	12	AATAAA	17
PC1244	#2162-91	after spacer	GCCTGCTCAG	12	AATAAA	17
PC1254	#2776-93	after spacer	GCCTGCTCAG	12	AATAAA	17
PC1203	#2528-90	after spacer	GCCTGCTCAG	12	AATAAA	17
PC1199	#CG575	after spacer	GCCTGCTCAG	12 (NFI")	N.D.	19
PC1222	#CG575	after spacer	GCCTGCTCAG	12 (NFI")	(AATAAA)	15
PC1206	#CG575	after spacer	GCCTGCTCAG	12 (NFI")	N.D.	16
PC1175	#275	after spacer	GCCTGCTCAG	12 (NFI")	(AATAAA)	16
PC1237	#2617-93	after spacer	GCCTGCTCAG	12 ?	AATAAA	15
PC1207	#2540-87	after spacer	GCCTGCTCAG	12 ?	AATAAA	16
PC1239	#2344-94	after spacer	GCCTGCTCAG	12	AATAAA	18
PC1116	LL-14/SV40	after spacer	GCCTGCTCAG	12	AATAAA	18
PC1184	Myo163.1	after spacer	GCCTGCTCAG	12	(AATAGN)	17
PC1208	#2540-87	after spacer	GCCTGCTCAG	12	AATAAA	19
PC1224	#2540-87	after spacer	GCCTGCTCAG	12	AATAAA	47
PC1201	#837-88	after spacer	GCCTGCTCAG	12	AATAAA	15
PC1236	#2162-91	after spacer	GCCTGCTCAG	12	AATAAA	15
PC1243	#2162-91	after spacer	GCCTGCTCAG	12	AATAAA	28
PC1227	#183-89	after spacer	GCCTGCTCAG	12	?	7
			GCCTGCTCAG	12	(AATAGN)	14
3'-untranslated region (various positions)						
PC1193	#58	within 3'-UTR	TATCTTTTCA	12	AATAAA	23
PC1194	#58	within 3'-UTR	TATCTTTTCA	12	AATAAA	34
PC1195	#192	within 3'-UTR	TATCTTTTCA	12	AATAAA	23
PC1196	#192	within 3'-UTR	TATCTTTTCA	12	AATAAA	23
PC1189	Myo192.1	within 3'-UTR	TCTTCCCTCT	12	AATAAA	>17
			ATACACCTTA	12	AATAAA	23
PC1117	Ad-312/SV40	within 3'-UTR	TTCGCATGTT	12	AATAAA	23
PC1253	#2617-93	within 3'-UTR	CACCTTCATC	12	?	?
PC1264	#568-92	within 3'-UTR	ATAAGGACTA	12	N.D.	117-AB
PC1270	#2528-90	within 3'-UTR	NCCTGTNAGC	12	(ACTAAA)	19
			TAGAGATTAG	12	N.D.	4

N.T.: NOT TESTABLE.

N.T.<sup>1</sup>: LENGTH OF ECTOPIC SEQUENCE DOES NOT ALLOW DEVELOPMENT OF PRIMER-SETN.T.<sup>2</sup>: ECTOPIC SEQUENCE IS MAINLY COMPOSED OF REPETITIVE SEQUENCES

N.D.: NOT DETECTED

## LEGENDS TO THE FIGURES

## Figure 1

Long range physical map of a 6 Mb region on the long arm of human chromosome 12 deduced from a YAC contig consisting of 75 overlapping CEPH YAC clones and spanning the chromosome 12q breakpoints as present in a variety of benign solid tumors. The long range physical map of the composite genomic DNA covered by the YAC inserts is represented by a black solid line with the relative positions of the various restriction sites of rare cutting enzymes indicated. DNA regions in which additional cutting sites of a particular restriction enzyme might be found are indicated by arrows. Polymorphic restriction endonuclease sites are marked with asterisks. DNA markers isolated and defined by others are depicted in green. DNA markers obtained by us are shown in boxes and are labelled by an acronym (see also Table I and II). The relative positions of these DNA markers in the long range physical map are indicated and those corresponding to particular YAC ends are linked to these by a dotted line. Some of the DNA markers have been assigned to a DNA interval and this is indicated by arrows. For DNA markers in white boxes STSs have been developed and primer sets are given in Table II. For those in yellow boxes, no primer sets were developed. The DNA intervals containing RAP1B, EST01096, or IFNG are indicated. Where applicable, D number assignments are indicated. Below the long range physical map, the sizes and relative positions of the overlapping YAC clones fitting within the consensus long range restriction map are given as solid blue lines. DNA regions of YAC inserts not fitting within the consensus long range restriction map are represented by dotted blue lines. CEPH microtiter plate addresses of the YAC clones are listed. The orientation of the YAC contig on chromosome 12 is given. The relative positions of ULCR12 and MAR are indicated by red solid lines labelled by the corresponding acronyms. Accession numbers of STSs not listed in Table I: CH9 (#U27142); RM1 (#U29049); RM110 (#U29022);

RM111 (#U29023); RM130 (#U27139); RM131 (#U29001); RM132 (#U27138); RM133 (#U27137). Restriction sites: B: BssHII; K: KspI (=SacII); M: MluI; N: NotI; P: PvuI; Sf: SfiI.

## 5 Figure 2

Contig of overlapping cosmids, long range restriction and STS map spanning a segment of MAR of about 445 kb. Contig elements are numbered and defined in the list below. LL12NC01-derived cosmid clones are named after their microtiter plate addresses. GenBank accession numbers (#) of the various STSs are listed below. STSs are given in abbreviated form; e.g. RM33 instead of STS 12-RM33. A 40 kb gap between STSs "K" and "O" in the cosmid contig was covered by  $\lambda$  clones (clones 38 and 40) and PCR products (clones 37 and 39). The orientation of the contig on the long arm of chromosome 12 is given as well as the order of 37 STSs (indicated in boxes or labelled with encircled capital letters). The slanted lines and arrows around some of the STS symbols at the top of the figure mark the region to which the particular STS has been assigned. It should be noted that the cosmid contig is not scaled; black squares indicate STSs of cosmid ends whereas the presence of STSs corresponding to internal cosmid sequences are represented by dots. Long range restriction map: Bs: BssHII; K: KspI (=SacII); M: MluI; N: NotI; P: PvuI; Sf: SfiI. At the bottom of the figure, detailed restriction maps are shown of those regions containing exons (boxes below) of the HMGI-C gene. Noncoding sequences are represented by open boxes and coding sequences by black boxes. Estimated sizes (kb) of introns are as indicated. The relative positions of the translation initiation (ATG) and stop (TAG) codons in the HMGI-C gene as well as the putative poly-adenylation signal are indicated by arrows. Detailed restriction map: B: BamHI; E: EcoRI; H: HindIII. MAR: Multiple Aberration Region; DBD: DNA Binding Domain.

1=140A3	11=142G8	21=124D8	31=59A1	41=128A2	51=65E6
2=202A1	12=154A10	22=128A7	32=101D8	42=142H1	52=196E1
3=78F11	13=163D1	23=129F9	33=175C7	43=204A10	53=215A8

4=80C9	14=42H7	24=181C1	34=185H2	44=145E1	54=147G8
5=109B12	15=113A5	25=238E1	35=189C2	45=245E8	55=211A9
6=148C12	16=191H5	26=69B1	36=154B12	46=154F9	56=22D8
7=14H6	17=248E4	27=260C7	37=pRM150	47=62D8	57=116B7
5. 8=51F8	18=33H7	28=156A4	38=pRM144	48=104A4	58=144D12
9=57C3	19=50D7	29=27E12	39=PKXL	49=184A9	
10=86A10	20=68B12	30=46G3	40=pRM147	50=56C2	
A = STS 12-EM12 (#U27145)	I = STS 12-CH12 (#U27153)				
10 Q = STS 12-RM120 (#U27161)	B = STS 12-EM30 (#U27146)				
J = STS 12-EM10 (#U27154)	R = STS 12-RM118 (#U27162)				
C = STS 12-EM14 (#U27147)	K = STS 12-EM37 (#U27155)				
S = STS 12-RM119 (#U27163)	D = STS 12-EM31 (#U27148)				
L = STS 12-RM146 (#U27156)	T = STS 12-EM2 (#U27164)				
15 E = STS 12-CH11 (#U27149)	M = STS 12-RM145 (#U27157)				
U = STS 12-EM4 (#U27165)	F = STS 12-EM18 (#U27150)				
N = STS 12-RM151 (#U27158)	V = STS 12-EM3 (#U27166)				
G = STS 12-EM11 (#U27151)	O = STS 12-EM16 (#U27159)				
W = STS 12-EM15 (#U27167)	H = STS 12-CH10 (#U27152)				
20 P = STS 12-EM1 (#U27160)	X = STS 12-EM17 (#U27168)				
STS 12-CH5 (#U27136)	STS 12-CH9 (#U27142)				
STS 12-RM33 (#U27131)	STS 12-RM53 (#U27134)				
STS 12-RM76 (#U27132)	STS 12-RM86 (#U27133)				
25 STS 12-RM98 (#U26647)	STS 12-RM99 (#U27130)				
STS 12-RM103 (#U26689)	STS 12-RM130 (#U27139)				
STS 12-RM132 (#U27138)	STS 12-RM133 (#U27137)				
STS 12-RM151 (#U27158)					

30

**Figure 3**

Schematic representation of FISH mapping data obtained for tumor cell lines with chromosome 12q13-q15 aberrations, including 8 lipoma, 10 uterine leiomyoma, and 8 pleomorphic salivary gland adenoma cell lines in consecutive experiments following our earlier FISH studies. Probes used included phage clones pRM144 (corresponding STSs: RM86 and RM130) and pRM147 (RM151), and cosmid clones 7D3 or 152F2 (RM103), 154F9 (CH9), 27E12 (EM11), 211A9 (RM33), 245E8 (RM53), 185H2 (RM76), 202A1 (RM98), 142H1 (RM99), 154B12 (RM132), and 124D8 (RM133). The DNA interval between RM33 and RM98 is estimated to be about 445 kb. Dots indicate conclusive FISH experiments that were performed on metaphase chromosomes of a particular cell line using as molecular

probe, a clone containing the STS given in the box above. Solid lines indicate DNA intervals to which a breakpoint of a particular cell line was concluded to be mapping. Open triangles indicate deletions observed during FISH analysis.

- 5 Open circles indicate results of FISH experiments on metaphase chromosomes of Li-501/SV40 cells with hybridization signals on a cytogenetically normal chromosome 3. The positions of chromosome 12 breakpoints of tumor cell lines mapping outside MAR are indicated by arrows. The
- 10 molecularly cloned breakpoints of LM-30.1/SV40 and LM-608/SV40 are indicated by asterisks. Breakpoints in various uterine leiomyoma cell lines splitting cosmid 27E12 (EM11) are indicated by "across".

15 **Figure 4**

3'-RACE product comprising the junction between part of the HMGI-C gene and part of the LPP gene. The primers used and the junction are indicated. The cDNA synthesis was internally primed and not on the true poly(A)

20 tail.

**Figure 5**

Partial cDNA sequence of the LPP gene.

25 **Figure 6**

Amino acid sequence of the LPP gene. LIM domains are boxed. The breaking point is indicated with an arrow.

**Figure 7**

Nucleotide sequence of HMGI-C (U28749). The transcription start site indicated as proposed by Manfioletti et al. [67] was arbitrarily chosen as a start 5' site. The sequence contains the complete coding sequence.

**Figure 8**

Gel of PCR products obtained as described in Example 5.

## REFERENCES

1. Sreekantaiah, C., Leong, S.L.P., Karakousis, C.P., McGee, D.L., Rappaport, W.D., Villar, H.V., Neal, D., Fleming, S., Wankel, A., Herrington, P.N., Carmona, R. and Sandberg, A.A. (1991). Cytogenetic profile of 109 lipomas. *Cancer Res.* 51: 422-433.
2. Nilbert, M. and Heim, S. (1990). Uterine leiomyoma cytogenetics. *Genes Chrom. Cancer* 2: 3-13.
3. Pandis, N., Heim, S., Willen, H., Bardi, G., Floderus, U.M., Mandahl, N. and Mitelman, F. (1991). Chromosome analysis of 96 uterine leiomyomas. *Cancer Genet. Cytogenet.* 55: 11-18.
4. Sandros, J., Stenman, G. and Mark, J. (1990). Cytogenetic and molecular observations in human and experimental salivary gland tumours. *Cancer Genet. Cytogenet.* 44: 153-167.
5. Bullerdiek, J., Wobst, G., Meyer-Bolte, K., Chilla, R., Haubrich, J., Thode, B. and Bartnitzke, S. (1993). Cytogenetic subtyping of 220 salivary gland pleomorphic adenomas: correlation to occurrence, histological subtype, and in vitro cellular behavior. *Cancer Genet. Cytogenet.* 65: 27-31.
6. Walter, T.A., Xuan Fan, S., Medchill, M.T., Berger, C.S., Decker, H-J.H. and Sandberg, A.A. (1989). Inv(12)(p11.2q13) in an endometrial polyp. *Cancer Genet. Cytogenet.* 41: 99-103.
7. Vanni, R., Dal Cin, P., Marras, S., Moerman, P., Andrtia, A., Valdes, E., Deprest, J., and Van den Berghe, H., (1993). Endometrial polyp: Another benign tumor characterized by 12q13-q15 changes. *Cancer Genet. Cytogenet.* 68: 32-33.
8. Mandahl, N., Orndal, C., Heim, S., Willen, H., Rydholm, A., Bauer, H.C.F. and Mitelman, F. (1993). Aberrations of chromosome segment 12q13-15 characterize a subgroup of hemangiopericytomas. *Cancer* 71: 3009-3013.
9. Mandahl, N., Heim, S., Arheden, K., Rydholm, A., Willen, H. and Mitelman, F. (1989). Chromosomal rearrangements in chondromatous tumors. *Cancer* 65:242-248.
10. Bridge, J.A., Persons, D.L., Neff, J.R. and Bhatia, P. (1992). Clonal karyotypic aberrations in enchondroma. *Cancer Detect. Prev.* 16: 215-219.
11. Hirabayashi, Y., Yoshida, M.A., Ikeuchi, T., Ishida, T., Kojima, T. Higaki, S., Machinami, R. and Tonomura,

- A. (1992). Chromosome rearrangements at 12q13 in two cases of chondrosarcomas. *Cancer Genet. Cytogenet.* 60: 35-40.
12. Mandahl, N., Willen, H., Rydholm, A. and Mitelman, F. (1993). Rearrangement of band q13 on both chromosomes 12 in a periosteal chondroma. *Genes Chrom. Cancer* 6: 121-123.
  13. Dal Cin, P., Kools, P., De Jonge, I., Moerman, Ph., Van de Ven W., Van den Berghe H. (1993a). Rearrangement of 12q14-15 in Pulmonary Chondroid Hamartoma. *Genes Chrom. Cancer*, 8, 131-133.
  14. Schoenberg Fejzo, M., Yoon, S.J., Montgomery, K.T., Rein, M.S., Weremowicz, S., Krauter, K.S., Dorman, T.E., Fletcher, J.A., Mao, J., Moir, D.T., Kucherlapati, R.S., and Morton, C.C. (1995). Identification of a YAC spanning the translocation breakpoints in uterine leiomyomata, pulmonary chondroid hamartoma and lipoma. Physical mapping of the 12q14-15 breakpoint region in uterine leiomyomata. *Genomics* 26: 265-275.
  15. Birdsall, S.H., MacLennan, K.A. and Gusterson, B.A. (1992). t(6;12)(q23;q13) and t(10;16)(q22;p11) in a phyllodes tumor of the breast. *Cancer Genet. Cytogenet.* 60: 74-77.
  16. Rohen, C., Bonk, U., Staats, B., Bartnizke, S. and Bullerdiek, J. (1993). Two human breast tumors with translocations involving 12q13-15 as the sole cytogenetic abnormality. *Cancer Genet. Cytogenet.*, 69: 68-71.
  17. Jenkins, R.B., Kimmell, D.W., Moertel, C.A., Schulz, C.A., Menezes, R.M., Scheihauer, B., Kelly, P.J. and Dewald, G.W. (1989). Recurrent cytogenetic abnormalities in 80 gliomas. *Cytogenet. Cell Genet.* 51: 1019.
  18. Noguera, R., Llombart-Bosch, A., Lopez-Gines, C., Carda, C. and Fernandez, Cl. (1989). Giant-cell tumor of bone, stage II, displaying translocation t(12;19)(q13;q13). *Virchows Archiv A Pathol. Anat.* 415: 377-382.
  19. Turc-Carel, C., Limon, J., Dal Cin, P., Rao, U., Karakousis, C., and Sandberg, A.A. (1986). Cytogenetic studies of adipose tissue tumours. II. Recurrent reciprocal translocation t(12;16)(q13;p11) in myxoid liposarcomas. *Cancer Genet. Cytogenet.* 23: 291-299.
  20. Rodriguez, E., Sreekantaiah, C., Reuter, V.E., Motzer, R.J. and Chaganti, R.S.K. (1992). t(12;22)(q13;q13) and trisomy 8 are nonrandom aberrations in clear-cell sarcoma. *Cancer Genet. Cytogenet.* 64: 107-110.

21. Reeves, B.R., Fletcher, C.D.M. and Gusterson, B.A. (1992). Translocation t(12;22)(q13;q13) is a nonrandom rearrangement in clear cell sarcoma. *Cancer Genet. Cytogenet.* 64: 101-103.
22. Fletcher, J.A. (1992). Translocation (12;22)(q13-14;q12) is a nonrandom aberration in soft-tissue clear-cell sarcoma. *Genes Chrom. Cancer* 5: 184.
23. Roberts, P., Browne, C.F., Lewis, I.J., Bailey, C.C., Spice, R.D., Williams, J. and Batcup, G. (1992). 12q13 Abnormality in rhabdomyosarcoma. A nonrandom Occurrence? *Cancer Genet. Cytogenet.* 60: 135-140.
24. Schoenmakers, H.F.P.M., Mols, R., Wanschura, S., Kools, P.F.J., Geurts, J.M.W., Bartnitzke, S., Bullerdiek, J., Van den Berghe, H., and Van de Ven, W.J.M. (1994b). Identification, molecular cloning and characterization of the chromosome 12 breakpoint cluster region of uterine leiomyomas. *Genes Chrom. & Cancer.* 11: 106-118.
25. Van de Ven, W.J.M., Schoenmakers, H.F.P.M., Wanschura, S., Kazmierczak, B., Kools, P.F.J., Geurts, J.M.W., Bartnitzke, S., Van den Berghe, H., and Bullerdiek, J. (1995). Molecular characterization of MAR, a multiple aberration region on human chromosome segment 12q13-q15 implicated in various solid tumors. *Genes Chrom. & Cancer.* 12: 296-303. (Enclosed as "ANNEX 1")
26. Casalone, r. et al. (1991) Cytogenetic analysis reveals clonal proliferation of smooth muscle cells in atherosclerotic plaques. *Hum. Gen.* 87, 139-143.
27. Vanni, R. et al. (1990) Atherosclerotic plaque as a benign tumor. *Cancer Genet. Cytogenet.* 47, 273-274.
28. Giancotti, V. et al. Elevated levels of a specific class of nuclear phosphoproteins in cells transformed with ras and v-mos oncogenes and by co-transfection with c-myc and polyoma middle T genes. *EMBO J.* 6, 1981-1987 (1987).
29. Warburton, D., Gersen, S., Yu, M-T., Jackson, C., Handelin, B. and Housman, D. (1990). Monochromosomal rodent-human hybrids from microcell fusion of human lymphoblastoid cells containing an inserted dominant selectable marker. *Genomics* 6: 358-366.
30. Schoenmakers, H.F.P.M., Kools, P.F.J., Kazmierczak, B., Bullerdiek, J., Claussen, U., Horsthemke, B., Van den Berghe, H. and Van de Ven, W.J.M. (1993). Isolation of a somatic cell hybrid retaining the der(16)t(12;16)(q13;p11.2) from a myxoid liposarcoma cell line. *Cytogenet. Cell Genet.* 62: 159-161.
31. Rychlik, W. and Rhoads, R.E. (1989). A computer program for choosing optimal oligonucleotides for filter

hybridization, sequencing and in vitro amplification of DNA. *Nucleic Acids Res.* 17: 8543-8551.

32. Verbeek, J.S., Roebroek, A.J.M., Van den Ouweland, A.M.W., Bloemers, H.P.J. and Van de Ven, W.J.M. (1985). Human c-fms proto-oncogene: comparative analysis with an abnormal allele. *Mol. Cell. Biol.* 5: 422-426.
33. Montgomery, K.T., LeBlanc, J.M., Tsai, P., McNinch, J.S., Ward, D.C., De Jong, P.J., Kucherlapati, R., and Krauter, K.S. (1993). Characterization of two chromosome 12 cosmid libraries and development STSs from cosmids mapped by FISH. *Genomics* 17: 682-693
34. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
35. Dal Cin, P., Kools, P., Scot, R., De Wever, I., Van Damme, B., Van de Ven, W. and Van den Berghe, H. (1993b). Molecular cytogenetic characterization of ring chromosomes in adipose tissue tumors. *Cancer Genet. Cytogenet.*, 68, 85-90.
36. Kievits, T., Dauwerse, J.G., Wiegant, J., Devilee, P., Breuning, M.H., Cornelisse, C.J. and van Ommen, G., Pearson, P.L. (1990). Rapid subchromosomal localization of cosmids by nonradioactive in situ hybridization. *Cytogenet. Cell Genet.* 53: 134-136.
37. Albertsen, H.M., Abderrahim, H., Cann, H.M., Dausset, J., L Paslier, D., and Cohen, D. (1990). Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. *Proc. Natl. Acad. Sci. USA* 87: 4256-4260.
38. Chumakov, I., Rigault, P., Guillou, S., Ougen, P., Billaut, A., Guasconi, G., Gervy, P., LeGall, I., Soularue, P., Grinas, L., Bougueleret, L., Bellanné-Chantelot, C., Lacroix, B., Barillot, E., Gesnouin, P., Pook, S., Vaysseix, G., Frelat, G., Schmitz, A., Sambucy, J.L., Bosch, A., Estivill, X., Weissenbach, J., Vignal, A., Riethman, H., Cox, D., Patterson, D., Gardiner, K., Hattori, M., Sakaki, Y., Ichikawa, H., Ohki, M., L Paslier, D., Heilig, R., Antonarakis, S., and Cohen, D. (1992). Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature* 359: 380-387.
39. Schoenmakers, H.F.P.M., Kools, P.F.J., Mols, R., Kazmierczak, B., Bartnitzke, S., Bullerdiek, J., Dal Cin, P., De Jong, P.J., Van den Berghe, H. and Van de Ven, W.J.M. (1994a). Physical mapping of chromosome 12q breakpoints in lipoma, pleomorphic salivary gland adenoma, uterine leiomyoma, and myxoid liposarcoma. *Genomics* 20: 210-222.

40. Chu, G., Vollrath, D. and Davis, R.W. (1986) Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science* 234: 1582-1585.
41. Geurts, J., Schoenmakers H.F.P.M., Mols, R., and Van de Ven, W.J.M. (1994). Improved procedure for rapid isolation and sequencing of DNA insert termini in yeast artificial chromosomes. *Meth. Mol. Cell. Biol.*, 4: 257-265.
42. Kools, P.F.J., Wanschura, S., Schoenmakers, E.F.P.M., Geurts, J.W.M., Mols, R., Kazmierczak, B., Bullerdiel, J., Van den Berghe, H. and Van de Ven, W.J.M. (1995). Identification of the chromosome 12 translocation breakpoint region of a pleomorphic salivary gland adenoma with t(1;12)(p22;q15) as the sole cytogenetic abnormality. *Cancer Genet. Cytogenet.*, 79: 1-7.  
(Enclosed as "ANNEX 2")
43. Kazmierczak, B., Wanschura, S., Rosigkeit, J., Meyer-Bolte, K., Uschinsky, K., Haupt, R., Schoenmakers, E.F.P.M., Bartnitzke, S., Van de Ven, W.J.M. and Bullerdiel, J. (1995). Molecular characterization of 12q14-15 rearrangements in three pulmonary chondroid hamartomas. *Cancer Res.*, in press.
44. Kucherlapati, R., Craig, I., and Marynen, P. (1994). Report of the second international workshop on human chromosome 12 mapping 1994. *Cytogenet. Cell Genet.* 67: 246-276
45. Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M. and Weissenbach, J. (1994). The 1993-94 Génethon human genetic linkage map. *Nature Genetics* 7: 246-339.
46. Ulinowski, Z., Taylor, K., Griffin, D., Delhanty, J. and Wolfe, J. (1991). D12S56: a highly polymorphic locus on human chromosome 12q14. *Ann. Hum. Genet.* 55: 279-282.
47. Trent, J.M., Olson, S. and Lawn, R.M. (1982) Chromosomal localization of human leukocyte, fibroblast, and immune interferon genes by means of in situ hybridization. *Proc. Natl. Acad. Sci. U.S.A.* 79: 7809-7813.
48. Pizon, V., Lerosey, I., Chardin, P. and Tavitian, A. (1988). Nucleotide sequence of a human cDNA encoding a ras-related protein (rap1B). *Nucleic Acids Res.* 16: 7719.
49. Adams, M.D., Dubnick, M., Kerlavage, A.R., Moreno, R., Kelley, J.M., Utterback, T.R., Nagle, J.W., Fields, C., and Venter, J.C. (1992). Sequence identification of 2,375 human brain genes. *Nature* 355: 632-634.

50. Cohen, D., Chumakov, I., and Weissenbach, J. (1993). A first-generation physical map of the human genome. *Nature* 366: 698-701.
51. Larsen, F., Gundersen, G., and Prydz, H. (1992a). Choice of enzymes for mapping based on CpG islands in the human genome. *GATA* 9: 80-85.
52. Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992b). CpG islands as gene markers in the human genome. *Genomics* 13: 1095-1107.
53. Lavia, P., MacLeod, D., and Bird, A. (1987). Coincident start sites for divergent transcripts at randomly selected CPG-rich island of mouse. *EMBO J.* 6: 2773-2779.
54. Kazmierczak, B., Bartnitzke, S., Hartl, M. & Bullerdiel, J. In vitro transformation by the SV40 "early region" of cells from a human benign salivary gland tumor with a 12q13-q15 rearrangement. *Cytogenet. Cell Genet.* 53, 37-39 (1990).
55. Albertsen, H.M., Adderrahim, H., Cann, H.M., Dausset, J., Le Paslier, D. & Cohen, D. Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. *Proc. natl. Acad. Sci. U.S.A.* 87, 4256-4260 (1990).
56. Chumakov, I. et al. Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature* 359, 380-387 (1992).
57. Green, E.D. & Olson, M.V. Systematic screening of yeast artificial-chromosome libraries using the polymerase chain reaction. *Proc. natl. Acad. Sci. U.S.A.* 87, 1213-1217 (1990).
58. Montgomery, K.T. et al. Characterization of two chromosome 12 cosmid libraries and development of STSS from cosmids mapped by FISH. *Genomics* 17, 682-693 (1993).
59. Geurts, J.M.W., Schoenmakers E.F.P.M., Mols, R. & Van de Ven, W.J.M. Improved procedure for rapid isolation and sequencing of DNA insert termini in yeast artificial chromosomes. *Meth. Mol. Cell. Biol.* 4, 257-265 (1994).
60. Nelson, D.L. et al. Alu polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources. *Proc. natl. Acad. Sci. U.S.A.* 86, 6686-6690 (1989).
61. Rychlik, W. & Rhoads, R.E. A computer program for choosing optimal oligonucleotides for filter

- hybridization, sequencing and in vitro amplification of DNA. *Nucl. Acids Res.* 17, 8543-8551 (1989).
62. Feinberg, A.P. & Vogelstein, B. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132, 6-13 (1984).
  63. Smith, M.W., Holmsen, A.L., Wei, Y.H., Peterson, M. & Evans, G.A. Genomic sequence sampling: a strategy for high resolution sequence-based physical mapping of complex genomes. *Nature Genetics* 7, 40-47 (1994).
  64. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410 (1990).
  65. Patel, U.A. et al. Expression and cDNA cloning of human HMGI-C phosphoprotein. *Biochem. Biophys. Res. Commun.* 201, 63-70 (1994).
  66. Bustin, M., Lehn, D.A. & Landsman, D. Structural features of the HMG chromosomal proteins and their genes. *Biochim. Biophys. Acta* 1049, 231-243 (1990).
  67. Manfioletti, G. et al. cDNA cloning of the HMGI-C phosphoprotein, a nuclear protein associated with neoplastic and undifferentiated phenotypes. *Nucl. Acids Res.* 19, 6793-6797 (1991).
  68. Wagner, M.J., Ge, Y., Siciliano, M. & Wells, D.E. A hybrid cell mapping panel for regional localization of probes to human chromosome 8. *Genomics* 10, 114-125 (1991).
  69. Friedman, M., Holth, L.T., Zoghbi, H.Y. & Reeves, R. Organization, inducible-expression and chromosome localization of the human HMG-I(Y) nonhistone protein gene. *Nucl. Acids Res.* 21, 4259-4267 (1993).
  70. Rabbitts, T.H. Chromosomal translocations in human cancer. *Nature* 372, 143-149 (1994).
  71. Yang-Yen, H.F. & Rothblum, L.I. Purification and characterization of a high-mobility-group-like DNA-binding protein that stimulates rRNA synthesis in vitro. *Mol. Cell. Biol.* 8, 3406-3414 (1988).
  72. Reeves, R., Langan, T.A. & Nissen, M.S. Phosphorylation of the DNA-binding domain of nonhistone high-mobility group I protein by cdc2 kinase: reduction of binding affinity. *Proc. natl. Acad. Sci. U.S.A.* 88, 1671-1675 (1991).
  73. Thanos, D. & Maniatis, T. The high mobility group protein HMGI(Y) is required for NF- $\kappa$ B-dependent virus

- induction of the human IFN- $\beta$  gene. Cell 71, 777-789 (1992).
74. Du, W. & Maniatis, T. The high mobility group protein HMGI(Y) can stimulate or inhibit DNA binding of distinct transcriptional factor ATF-2 isoforms. Proc. natl. Acad. Sci. U.S.A. 91, 11318-11322 (1994).
  75. Aizawa, S., Nishino, H., Saito, K., Kimura, K., Shirakawa, H. & Yoshida, M. Stimulation of transcription in cultured cells by high mobility group protein 1: Essential role of the acidic carboxy-terminal region. Biochemistry 33, 14690-14695 (1994).
  76. Mitelman F (1994): Catalog of Chromosome Aberrations in Cancer. 4th ed., New York, Wiley- Liss.
  77. Miles, J.H., Zonanna, J., Mcfarlane, J., Aleck, K.A. & Bawle, E. Macrocephaly with hamartomas: Bannayan-Zonana syndrome. Am. J. Med. Genet. 19, 225-234 (1984).
  78. Crawford, A.W., Pino, J.D. & Beckerie, M.C.J. Cell. Biol., 124, 117-127 (1994).
  79. Schmeichel, K.L. & Beckerie, M.C. Cell, 79, 211-219 (1994)
  80. Sanchez-Garcia, I. & Rabbitts, T. (1994) The LIM domain: a new structural motif found in zinc-finger-like proteins. Trends in Genetics 10(9), 315-320.
  81. Compton, J. (1991). Nucleic acid sequence-based amplification. Nature 350, 91-92.

## ANNEX 1

Genes, Chromosome & Cancer 12:296-303 (1995)

5 **Molecular Characterization of MAR, a Multiple Aberration  
Region on Human Chromosome Segment 12q13-q15 Implicated  
in Various Solid Tumors**

Wim J.M. Van de Ven, Eric F.P.M. Schoenmakers, Sylke  
10 Wanschura, Bernd Kazmierczak, Patrick F.J. Kools, Jan  
M.W. Geurts, Sabine Bartnitzke, Herman Van den Berghe,  
and Jörn Bullerdiek

Center for Human Genetics, University of Leuven, Belgium  
15 (W.J.M.V.D.V., E.F.P.M.S., P.F.J.K., J.W.M.G., H.V.D.B.);

Center for Human Genetics, University of Bremen, Germany  
(S.W., B.K., S.B., J.B.).

20

Chromosome arm 12q breakpoints in seven cell lines  
derived from primary pleomorphic salivary gland adenomas  
were mapped by FISH analysis relative to nine DNA probes.  
These probes all reside in a 2.8 Mb genomic DNA region of  
25 chromosome segment 12q13-q15 and correspond to previously  
published sequence-tagged sites (STS). Their relative  
positions were established on the basis of YAC cloning  
and long range physical and STS content mapping. The 12q  
breakpoints of five of the cell lines were found to be  
30 mapping within three different subregions of the 445 kb  
DNA interval that was recently defined as the uterine  
leiomyoma cluster region of chromosome 12 breakpoints  
(ULCR12) between STS RM33 and RM98. All seven breakpoints  
appeared to map within the 1.7 Mb DNA region between STS

35

Received September 27, 1994; accepted November 7, 1994. Address reprint requests to Dr. Wim  
J.M. van de Ven, Center for Human Genetics, University of Leuven, Herestraat 49, B-3000  
Leuven, Belgium.

RM36 and RM103. Furthermore, the chromosome 12 breakpoints of three primary pleomorphic salivary gland adenomas were also found to be mapping between RM36 and RM103. Finally, FISH analysis of two lipoma cell lines with 12q13-q15 aberrations pinpointed the breakpoints of these to relatively small and adjacent DNA segments which, as well as those of two primary lipomas, appeared to be located also between RM36 and RM103. We conclude from the observed clustering of the 12q breakpoints of the three distinct solid tumor types that the 1.7 Mb DNA region of the long arm of chromosome 12 between RM36 and RM103 is a multiple aberration region which we designate MAR. Genes Chromosom Cancer 12:296-303 (1995). © 1995 Wiley-Liss, Inc.

15

#### INTRODUCTION

Chromosome translocations involving region q13-q15 of chromosome 12 have been observed in a wide variety of solid tumors (Mitelman, 1991). In subgroups of cytogenetically abnormal uterine leiomyomas (Nilbert and Heim, 1990; Pandis et al., 1991), pleomorphic salivary gland adenomas (Sandrois et al., 1990; Bullerdiek et al., 1993), and benign adipose tissue tumors (Sreekantaiah et al., 1991), 12q13-q15 aberrations are frequently observed. In a recent study (Schoenmakers et al., 1994b), we identified and molecularly characterized ULCR12, the uterine leiomyoma cluster region of chromosome 12 breakpoints. In the present study, we focus on the chromosome arm 12q breakpoints in pleomorphic adenoma of the salivary glands, a benign epithelial tumor originating from the major or minor salivary glands. It is the most common type of salivary gland tumor and accounts for almost 50% of all neoplasms in these organs. About 85% of the tumors are found in the parotid gland, 10% in the minor salivary glands, and 5% in the submandibular gland (Seifert et al., 1986). Although many of these adenomas appear to have a normal karyotype,

cytogenetic studies have also revealed recurrent specific chromosome anomalies (Sandros et al., 1990; Bullerdiak et al., 1993). Besides chromosome 8 aberrations, often translocations with a breakpoint in 8q12 with, as the most common aberration, a t(3;8)(p21;q12), aberrations of chromosome 12, usually translocations involving 12q13-q15, are also frequent. Non-recurrent clonal abnormalities have also been described. The frequent involvement of region 12q13-q15 in distinct solid tumor types suggests that this chromosomal region harbors gene(s) that might be implicated in the evolution of these tumors. Molecular cloning of the chromosome 12 breakpoints of these tumors and characterization of the junction fragments may therefore lead to the identification of such gene(s).

On the basis of fluorescence in situ hybridization (FISH) data, we have previously reported that the chromosome 12 breakpoints in a number of cell lines derived from primary pleomorphic salivary gland adenomas (Kazmierczak et al., 1990; Schoenmakers et al., 1994a), are located on the long arm of chromosome 12 in the interval between loci D12S19 and D12S8 (Schoenmakers et al., 1994a). This DNA interval has been estimated to be about 7cM (Keats et al., 1989; Craig et al., 1993). The interval containing the chromosome 12 breakpoints of these tumor cells was narrowed further by showing that all breakpoints mapped distally to the CHOP gene, which is directly affected by the characteristic t(12;16) translocation in myxoid liposarcomas (Aman et al., 1992; Crozat et al., 1993; Rabbitts et al., 1993) and is located between D12S19 and D12S8. In more recent studies (Kools et al., 1995), the chromosome 12 breakpoint of pleomorphic salivary gland adenoma cell line Ad-312/SV40 was pinpointed to a DNA region between sequence-tagged sites (STSs) RM110 and RM111, which is less than 165 kb in size. FISH evaluation of the chromosome 12 breakpoints of the other pleomorphic salivary gland adenoma cell lines indicated that they must be located proximally to

the one in Ad-312/SV40, at a distance of more than 800 kb (Kools et al., 1995). These results pointed towards a possible dispersion of the chromosome 12 breakpoints over a relatively large genomic region on the long arm of chromosome 12.

Here, we report physical mapping of the chromosome 12 breakpoints in pleomorphic salivary gland adenoma cells from primary tumors as well as established tumor cell lines. The karyotypic anomalies observed in the cells were all different but always involved region q13-q15 of chromosome 12. Using DNA probes between D12S8 and CHOP, which corresponded to sequence-tagged sites (STSs) of a long-range physical map of a 6 Mb DNA region and were obtained during chromosome walking experiments, we performed FISH experiments and defined more precisely a major chromosome 12 breakpoint cluster region of pleomorphic salivary gland adenoma. This breakpoint cluster region appeared to overlap with ULCR12. Furthermore, we tested whether 12q13-q15 breakpoints of lipomas might also map within the same region as those of pleomorphic salivary gland adenoma and uterine leiomyoma.

#### MATERIALS AND METHODS

##### Primary Solid Tumors and Derivative Cell Lines.

Primary solid tumors including pleomorphic salivary gland adenomas, lipomas, and uterine leiomyomas were obtained from the University Clinics in Leuven, Belgium (Dr. I. De Wever); in Bremen, Germany (Dr. R. Chille); in Krefeld, Germany (Dr. J. Haubrich); and from the Institute of Pathology in Göteborg, Sweden (Dr. G. Stenman). For cell culturing and subsequent FISH analysis, tumor samples were finely minced, treated for 4-6 hours with 0.8% collagenase (Boehringer, Mannheim, FRG), and processed further for FISH analysis according to routine procedures.

Human tumor cell lines used in this study included the previously described pleomorphic salivary

gland adenoma cell lines Ad-211/SV40, Ad-248/SV40, Ad-263/SV40, Ad-295/SV40, Ad-302/SV40, AD-366/SV40, and Ad-386/SV40 (Kazmierczak et al., 1990; Schoenmakers et al., 1994a) and the lipoma cell lines Li-14/SV40 (Schoenmakers  
5 et al., 1994a) and recently developed Li-538/SV40.

Chromosome 12 aberrations found in these cell lines are listed in Table 1. Cells were propagated in TC199 culture medium with Earle's salts supplemented with 20% fetal bovine serum.

TABLE 1. Chromosome 12 Aberrations in Primary Human Solid Tumors and Cell Lines\*

Aberration	
5	-----
Cell lines	
Ad-211/SV40	t(8;12)(q21;q13-q15)
Ad-248/SV40	ins(12;6)(q15;q16q21)
Ad-263/SV40	inv(12)(q15q24.1)
10 Ad-295/SV40	t(8;12;18)(p12;q14;p11.2)
Ad-302/SV40	t(7;12)(q31;q14)
Ad-366/SV40	inv(12)(p13q15)
Ad-386/SV40	t(12;14)(q13-q15;q13-q15)
15 Li-14/SV40	t(3;12)(q28;q13)
Li-538/SV40	t(3;12)(q27;q14)
LM-5.1/SV40	t(12;15)(q15;q24)
LM-30.1/SV40	t(12;14)(q15;q24)
20 LM-65/SV40	t(12;14)(q15;q24)
LM-67/SV40	t(12;14)(q13-q15;q24)
LM-100/SV40	t(12;14)(q15;q24)
LM-605/SV40	ins(12;11)(q14;q21qter)
LM-608/SV40	t(12;14)(q15;q24)
25 LM-609/SV40	t(12;14)(q15;q24)
Primary tumors	
Ad-386	t(12;14)(q15;q11.2)
Ad-396	t(3;12)
30 Ad-400	t(12;16)
Li-166	t(12;12)
Li-167	t(3;12)(q28;q14-q15)
35 LM-163.1	t(12;14)(q14;q24)
LM-163.2	t(12;14)(q14-q24)
LM-168.3	t(X;12)(q22;q15)
LM-192	t(2;3;12)(q35;p21;q14)
LM-196.4	t(12;14)(q14;q24)
40	-----

\*Ad, pleomorphic salivary gland adenoma; Li, lipoma; uterine leiomyoma.

#### DNA probes.

In the context of a human genome project focusing on the long arm of chromosome 12, we isolated cosmid clones cRM33, cRM36, cRM51, cRM69, cRM72, cRM76, 5 cRM98, cRM103, and cRM133, from chromosome 12-specific arrayed cosmid library LLNL12NC01 (Montgomery et al., 1993). Further details of these cosmid clones have been reported at the Second International Chromosome 12 Workshop (1994) and will be described elsewhere 10 (Kucherlapati et al., 1994). Briefly, initial screenings were performed using a PCR-based screening strategy (Green and Olson, 1990), followed by filter hybridization analysis as the final screening step, as previously described (Schoenmakers et al., 1994b). The cosmid clones 15 were isolated using STSS derived from YAC clones. STSS were obtained upon rescue of YAC insert-ends using a methodology involving vectorette-PCR followed by direct solid phase fluorescent sequencing of the PCR products (Geurts et al., 1994) or from inter-Alu PCR (Nelson et 20 al., 1989). Cosmid clones were grown and handled according to standard procedures (Sambrook et al., 1989).

Cosmid clone cPK12qter, which maps to the telomeric region of the long arm of chromosome 12 (Kools 25 et al., 1995) was used as a reference marker.

#### Chromosome Preparations and Fluorescence In Situ Hybridization.

Metaphase spreads of the pleomorphic salivary 30 gland adenoma cell lines or normal human lymphocytes were prepared as described before (Schoenmakers et al., 1993). To unambiguously establish the identity of chromosomes in the FISH experiments, FISH analysis was performed after GTG-banding of the same metaphase spreads. GTG- 35 banding was performed essentially as described by Smit et al. (1990). In situ hybridizations were carried out according to a protocol described by Kievits et al. (1990) with some minor modifications (Kools et al., 1994;

Schoenmakers et al., 1994b). Cosmid and YAC DNA was labelled with biotin-11-dUTP (Boehringer Mannheim) or biotin-14-dATP (BRL, Gaithersburg) as described before (Schoenmakers et al., 1994b). Specimens were analyzed on  
5 a Zeiss Axiophot fluorescence microscope using a FITC filter (Zeiss). Results were recorded on Scotch (3M) 640 asa film.

## RESULTS

10

### FISH Mapping of 12q Breakpoints in Cell Lines of Pleomorphic Salivary Gland Adenoma.

In previous studies (Schoenmakers et al., 1994a), we mapped the chromosome 12 breakpoints in a  
15 number of pleomorphic adenomas of the salivary glands relative to various DNA markers and established that these were all located proximally to locus D12S8 and distal to the CHOP gene. This region is somewhat smaller than the 7 cM region encompassed by linkage loci D12S8  
20 and D12S19 (Keats et al., 1989). Using YAC cloning, a long range physical/STS map has been constructed covering most of that 7 cM region, as recently reported (Kucherlapati et al., 1994). Furthermore, numerous genomic clones (cosmid clones) have been isolated and  
25 their relative positions within this map established (Kucherlapati et al., 1994). Nine of these cosmids, including CRM33, CRM36, CRM51, CRM69, CRM72, CRM76, CRM98, CRM103, and CRM133, were used in FISH studies to establish the positions of the chromosome 12 breakpoints  
30 of the seven cell lines derived from pleomorphic adenomas of the salivary glands (Table 1). The relative mapping order of these nine cosmid clones, which cover a genomic region on the long arm of chromosome 12 of about 2.8 Mb, is indicated in Figure 1 and the results of FISH studies  
35 with the various cosmid probes are schematically summarized in the same figure. As an illustration, FISH results obtained with metaphase cells of cell line Ad-295/SV40 using CRM76 and CRM103 as probes are shown in

Figure 2. It should be noted that for the identification of chromosomes, pre-FISH GTG-banding was used routinely. On the basis of such banding, hybridization signals could be assigned conclusively to chromosomes of known  
5 identity; this was of major importance for cases with cross- or background hybridization signals, as these were occasionally observed. When GTG-banding in combination with FISH analysis provided inconclusive results, either because of weak hybridization signals or rather vague  
10 banding, FISH experiments were performed with cosmid clone cPK12qter (Kools et al., 1995) as a reference probe.

FISH analysis of metaphase chromosomes of each of the seven pleomorphic salivary gland adenoma cell  
15 lines with cosmid cRM103 revealed that this cosmid mapped distal to the chromosome 12 breakpoints of all seven cell lines studied here. Metaphase chromosomes of six of the seven cell lines were also tested with probe cRM69 and, in two cases, with cRM51. The results of the latter  
20 experiments were always consistent with those obtained with cRM103. Similar FISH analysis with cRM36 as probe indicated that this probe mapped proximal to all the breakpoints. These results were always consistent with those obtained for five of the seven cell lines in  
25 experiments using cRM72. Altogether, the results of our FISH studies indicated that the chromosome 12 breakpoints of all seven cell lines map between cRM36 and cRM103, which spans a genomic region of about 1.7 Mb.

**Fine Mapping of 12q Breakpoints in Cell Lines Derived from Pleomorphic Adenomas of the Salivary Glands.**

For subsequent fine mapping of the chromosome 12 breakpoints of the seven pleomorphic salivary gland adenoma cell lines, additional FISH studies were performed, as schematically summarized in Figure 1. The breakpoints of cell lines Ad-211/SV40, Ad-295/SV40, and Ad-366/SV40 appeared to be located in the DNA region between CRM76 and CRM133, which was estimated to be about 75 kb. The breakpoints of the four other cell lines were found in different areas of the 1.7 Mb region between CRM36 and CRM103. That of cell line Ad-248/SV40 in a DNA segment of about 270 kb between CRM33 and CRM76, that of Ad-263/SV40 in a DNA segment of about 1 Mb between CRM98 and CRM103, that of Ad-302/SV40 in a DNA segment of about 240 kb between CRM33 and CRM36, and that of Ad-386/SV40 in a DNA segment of about 100 kb between CRM98 and CRM133. In conclusion, these results indicated that the chromosome 12 breakpoints of most (5 out of 7) of the cell lines are dispersed over the 445 kb genomic region on the long arm of chromosome 12 between CRM33 and CRM98. It is important to note already here that precisely this region was recently shown to contain the chromosome 12q breakpoints in cell lines derived from primary uterine leiomyomas (see Fig. 3) and was therefore designated ULCR12 (Schoenmakers et al., 1994b). As this segment of the long arm of chromosome 12 is involved in at least two types of solid tumors (Schoenmakers et al., 1994b; this study) and, as we will show below, also in a third solid tumor type, we will from now on refer to the DNA interval between CRM36 and CRM103 as MAR (multiple aberration region).

**FISH Mapping of 12q Breakpoints in Primary Pleomorphic Salivary Gland Adenomas.**

Our FISH studies on metaphase chromosomes of pleomorphic adenomas of the salivary glands presented so far were restricted to cell lines derived from primary

tumors. Although it is reasonable to assume that the chromosome 12 breakpoints in cell lines are similar if not identical to the ones in the corresponding primary tumors, differences as a result of the establishment of cell lines or subsequent cell culturing cannot fully be excluded. Therefore, we have investigated whether the chromosome 12 breakpoints in three primary salivary gland adenomas were mapping to MAR as well. To test this possibility, a combination of cosmid clones CRM33 and CRM103 were used as molecular probe. In all three cases, this cosmid pool clearly spanned the chromosome 12 breakpoints (data not shown), indicating that these breakpoints were indeed localized within MAR. In a recent study (Wanschura et al., submitted for publication), it was reported that the chromosome 12 breakpoints of five primary uterine leiomyomas with 12q14-15 aberrations were all found to cluster within the 1.5 Mb DNA fragment (between CRM33 and CRM103), which is known to harbor the breakpoints of various cell lines derived from primary uterine leiomyomas (schematically summarized in Fig. 3). Consistent with the results of the breakpoint mapping studies using cell lines, the results with the two primary solid tumor types establish that the breakpoints of the primary tumor cells are located in MAR.

#### **Chromosome Segment 12q13-q15 Breakpoints of Lipomas Mapping within MAR.**

To test the possibility that the chromosome 12 breakpoints of other solid tumors with 12q13-q15 aberrations also mapped within MAR, we studied two lipomas cell lines by FISH analysis---Li-14/SV40 and Li-538/SV40. The chromosome 12 aberrations of these two lipoma cell lines are given in Table 1. As molecular probes, cosmid clones CRM33, CRM53, CRM72, CRM76, CRM99, CRM103, and CRM133 were used. The breakpoint of Li-14/SV40 was mapped to the 75 kb DNA interval between RM76 and RM133, and that of Li-538/SV40 to the 90 kbp interval between RM76 and RM99 (data not shown), as schematically

illustrated in Figure 3. Similar FISH analysis of two primary lipomas using a mixture of CRM36 and CRM103 as molecular probe resulted in a hybridization pattern indicating that the mixture of probes detected sequences on either side of the breakpoints. These results are the first indications that also in lipoma, chromosome 12q13-q15 breakpoints occur that map within MAR. More lipoma cases should be tested to allow proper interpretation of this observation.

10

### DISCUSSION

In this study, we have mapped the chromosome 12 breakpoints of three primary pleomorphic salivary gland adenomas as well as seven established cell lines derived from such tumors. All breakpoints appeared to be located in a previously molecularly cloned and characterized chromosome DNA segment on the long arm of chromosome 12, of about 1.7 Mb in size, with five of them clustering in a DNA interval of less than 500 kb. The 1.7 Mb DNA region apparently contains a major breakpoint cluster region for this type of tumor. In a previous study, we have described the characterization of the chromosome 12 breakpoint of pleomorphic salivary gland adenoma cell line Ad-312/SV40 (Kools et al., 1995). The breakpoint of this cell line is now known to map at a distance of more than 2 Mb distally to this major breakpoint cluster region reported here. It is possible that the Ad-312/SV40 breakpoint involves other pathogenetically relevant genetic sequences than those affected by the clustered breakpoints. However, the possibility should not yet be excluded that all the 12q13-q15 breakpoints in pleomorphic salivary gland adenomas mapped so far belong to the same category and are dispersed over a relatively large DNA region of this chromosome, reminiscent of the 11q13 breakpoints in B-cell malignancies (Raynaud et al., 1993). More precise pinpointing of the various breakpoints could shed more light on this matter.

35

Of importance is the observation that the DNA segment that harbors the clustered 12q breakpoints of pleomorphic salivary gland adenomas appears to coincide with the DNA region that was recently defined as the uterine leiomyoma cluster region of chromosome 12 breakpoints, known as ULCR12 (Schoenmakers et al., 1994b). Of further interest is the fact that this region of chromosome 12 also harbors breakpoints of primary lipomas and lipoma cell lines derived from primary tumors with 12q13-q15 aberrations. Altogether, the results of all these studies now clearly demonstrate that chromosome 12 breakpoints of three distinct solid tumor types map to the same 1.7 Mb genomic region on the long arm of chromosome 12, establishing this region to be a multiple aberration region. To reflect this characteristic, we have designated this DNA segment MAR.

Genetic aberrations involving chromosomal region 12q13-q15 have been implicated by many cytogenetic studies in a variety of solid tumors other than the three already mentioned. Involvement of 12q13-q15 has also been reported for endometrial polyps (Walter et al., 1989; Vanni et al., 1993), clear cell sarcomas characterized by recurrent t(12;22)(q13;q13) (Fletcher, 1992; Reeves et al., 1992; Rodriguez et al., 1992), a subgroup of rhabdomyosarcoma (Roberts et al., 1992) and hemangiopericytoma (Mandahl et al., 1993a), chondromatous tumors (Mandahl et al., 1989; Bridge et al., 1992; Hirabayashi et al., 1992; Mandahl et al., 1993b), and hamartoma of the lung (Dal Cin et al., 1993). Finally, several case reports of solid tumors with involvement of chromosome region 12q13-q15 have been published---e.g., tumors of the breast (Birdsal et al., 1992; Rohen et al., 1993), diffuse astrocytomas (Jenkins et al., 1989), and a giant-cell tumor of the bone (Noguera et al., 1989). On the basis of results of cytogenetic studies, no predictions could be made about the relative distribution of the breakpoints of these tumor types. In light of the results of the present study, it would be of interest to

see whether the breakpoints of any of these solid tumors also map within or close to MAR. The various cosmid clones available now provide the means to test this readily.

5           The observation that 12q breakpoints of at least three different types of solid tumors map to the same DNA region is intriguing as it could be pointing towards the possibility that the same genetic sequences in MAR are pathogenetically relevant for tumor  
10 development in different tissues. If so, it is tempting to speculate that the gene(s) affected by the genetic aberrations might be involved in growth regulation. On the other hand, one cannot yet exclude the possibility that genetic sequences in MAR are not pathogenetically  
15 relevant, as the observed clustering of genetic aberrations in MAR could simply reflect genetic instability of this region, which becomes apparent in various solid tumors. To obtain more insight in this matter, the genes residing in MAR should be identified  
20 and characterized, and this can be achieved by various approaches using several techniques (Parrish and Nelson, 1993).

#### ACKNOWLEDGMENTS

25

          The constructive support of managing director G. Everaerts is greatly acknowledged. The authors would like to thank P. Dal Cin, J. Haubrich, R. Hille, G. Stenman, and I. De Wever for providing the solid tumor  
30 specimens studied in the present report; C. Huysmans, E. Meyen, K. Meyer-Bolte, R. Mols, and M. Willems for excellent technical assistance; and M. Leys for artwork. This work was supported in part by the EC through Biomed 1 program "Molecular Cytogenetics of Solid Tumours", the  
35 "Geconcerteerde Onderzoekacties 1992-1996", the National Fund for Scientific Research (NFWO; Kom op tegen Kanker), the "ASLK-programma voor Kankeronderzoek", the "Schwerpunktprogramm: Molekulare und Klassische

Tumorcytogenetik" of the Deutsche Forschungsgemeinschaft, and the Tönjes-Vagt Stiftung. This text presents results of the Belgian programme on Interuniversity Poles of attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by its authors. J.W.M. Geurts is an "Aspirant" of the National Fund for Scientific Research (NFWO; Kom op tegen Kanker).

## REFERENCES

- Aman P, Ron D, Mandahl N, Fioretos T, Heim S, Arheden K, Willen H, Rydholm A, Mitelman F (1992) Rearrangement of the transcription factor gene CHOP in myxoid liposarcomas with t(12;16)(q13;p11). Genes Chromosom Cancer 5: 278-285.
- Birdsal SH, MacLennan KA, Gusterson, BA (1992) t(6;12)(q23;q13) and t(10;16)(q22;p11) in a phyllodes tumor of the breast. Cancer Genet Cytogenet 60: 74-77.
- Bridge JA, Persons DL, Neff JR, Bhatia P (1992) Clonal karyotypic aberrations in enchondroma. Cancer Detect Prev 16: 215-219.
- Bullerdiel J, Wobst G, Meyer-Bolte K, Chilla R, Haubrich J, Thode B, Bartnitzke S (1993): Cytogenetic subtyping of 220 salivary gland pleomorphic adenomas: correlation to occurrence, histological subtype, and in vitro cellular behavior. Cancer Genet Cytogenet 65: 27-31.
- Craig IW, Gedde-Dahl T, Gemmill R, Kucherlapati R (1993) Report of the committee on the genetic constitution of chromosome 12. Genome Prior Rep 1: 402-418.
- Crozat A, Aman P, Mandahl N, Ron D (1993) Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. Nature 363: 640-644.
- Dal Cin P, Kools P, De Jonge I, Moerman Ph, Van de Ven W, Van den Berghe H (1993) Rearrangement of 12q14-15 in pulmonary chondroid hamartoma. Genes Chromosom Cancer 8: 131-133.
- Fletcher JA (1992) Translocation (12;22)(q13-14;q12) is a nonrandom aberration in soft-tissue clear-cell sarcoma. Genes Chromosom Cancer 5: 184.

Geurts JMW, Schoenmakers HFP, Mols R, Van de Ven WJM (1994) An improved procedure to quickly isolate and sequence the termini of DNA inserts of yeast artificial chromosomes. *Meth Mol Cell Biol* 4: 257-265.

Green ED, Olson MV (1990) Systematic screening of yeast artificial-chromosome libraries using the polymerase chain reaction. *Proc Natl Acad Sci USA* 87: 1213-1217.

Hirabayashi Y, Yoshida MA, Ikeuchi T, Ishida T, Kojima T, Higaki S, Machinami R, Tonomura A (1992) Chromosome rearrangements at 12q13 in two cases of chondrosarcomas. *Cancer Genet Cytogenet* 60: 35-40.

Jenkins RB, Kimmell DW, Moertel CA, Schulz CA, Menezes RM, Scheihauer B, Kelly PJ, Dewald GW (1989) Recurrent cytogenetic abnormalities in 80 gliomas. *Cytogenet Cell Genet* 51: 1019.

Kazmierczak B, Bartnitzke S, Hartl M, Bullerdiel J (1990): In vitro transformation by the SV40 "early region" of cells from a human benign salivary gland tumour with a 12q13-q15 rearrangement. *Cytogenet Cell Genet* 53: 37-39.

Keats B, Ott J, Conneally M (1989) Report of the committee on linkage and gene order. *Cytogenet Cell Genet* 51: 459-502.

Kievits T, Dauwerse JG, Wiegant J, Devilee P, Breuning MH, Cornelisse CJ, van Ommen G, Pearson PL (1990) Rapid subchromosomal localization of cosmids by nonradioactive in situ hybridization. *Cytogenet Cell Genet* 53: 134-136.

Kools PFJ, Roebroek AJM, Van de Velde HJK, Marynen P, Bullerdiel J, Van de Ven WJM (1993): Regional mapping of the human NSP gene to chromosome 14q21-q22 by

fluorescence in situ hybridization. Cytogenet Cell Genet 66:48-50.

Kools PFJ, Wanschura S, Schoenmakers EFPM, Geurts JWM, Mols R, Kazmierczak B, Bullerdiel J, Van den Berghe H, Van de Ven WJM (1995) Identification of the chromosome 12 translocation breakpoint region of a pleomorphic salivary gland adenoma with t(1;12)(p22;q15) as the sole cytogenetic abnormality. Cancer Genet Cytogenet, In press, 1995.

Kucherlapati R, Craig I, Marynen P (1994) Report of the second international workshop on human chromosome 12 mapping 1994. Cytogenet Cell Genet 67: 245-276.

Mandahl N, Heim S, Arheden K, Rydholm A, Willen H, Mitelman F (1989) Chromosomal rearrangements in chondromatous tumors. Cancer 65: 242-248.

Mandahl N, Orndal C, Heim S, Willen H, Rydholm A, Bauer HCF, Mitelman F (1993a) Aberrations of chromosome segment 12q13-15 characterize a subgroup of hemangiopericytomas. Cancer 71: 3009-3013.

Mandahl N, Willen H, Rydholm A, Mitelman F (1993b) Rearrangement of band q13 on both chromosomes 12 in a periosteal chondroma. Genes Chrom Cancer 6: 121-123.

Mitelman F (1991): Catalog of Chromosome Aberrations in Cancer. 4th ed., New York, Alan R. Liss.

Montgomery KT, LeBlanc JM, Tsai P, McNinch JS, Ward DC, De Jong PJ, Kucherlapati R, Krauter KS (1993) Characterization of two chromosome 12 cosmid libraries and development of STSs from cosmids mapped by FISH. Genomics 17:682-693.

Nelson DL, Ledbetter SA, Corbo L, Victoria MF, Ramirez-Solis R, Webster TD, Ledbetter DH, Caskey CT (1989) Alu polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources. Proc Natl Acad Sci USA 86: 6686-6690.

Nilbert M, Heim S (1990) Uterine leiomyoma cytogenetics. Genes Chromosom Cancer 2:3-13.

Noguera R, Llombart-Bosch A, Lopez-Gines C, Carda C, Fernandez Cl (1989) Giant-cell tumor of bone, stage II, displaying translocation t(12;19)(q13;q13). Virchows Archiv A Pathol Anat 415: 377-382.

Pandis N, Heim S, Bardi G, Flodérus U-M, Willén H, Mandahl N, Mitelman F (1991) Chromosome analysis of 96 uterine leiomyomas. Cancer Genet Cytogenet 55:11-18.

Parrish JE, Nelson DL (1993) Methods for finding genes a major rate-limiting step in positional cloning. GAT 10: 29-41.

Rabbitts TH, Forster A, Larson R, Nathan P (1993) Fusion of the dominant negative transcription regulator CHOP with a novel gene FUS by translocation t(12;16) in malignant liposarcoma. Nature Genet 4: 175-180.

Raynaud SD, Bekri S, Leroux D, Grosgeorge J, Klein B, Bastard C, Gaudray P, Simon MP (1993) Expanded range of 11q13 breakpoints with differing patterns of cyclin D1 expression in B-cell malignancies. Genes Chromosom Cancer 8: 80-87.

Reeves BR, Fletcher CDM, Gusterson BA (1992) Translocation t(12;22)(q13;q13) is a nonrandom rearrangement in clear cell sarcoma. Cancer Genet Cytogenet 64: 101-103.

Roberts P, Browne CF, Lewis IJ, Bailey CC, Spice RD, Williams J, Batcup G (1992) 12q13 Abnormality in rhabdomyosarcoma. A nonrandom Occurrence? Cancer Genet Cytogenet 60: 135-140.

Rodriguez E, Sreekantaiah C, Reuter VE, Motzer RJ, Chaganti RSK (1992) t(12;22)(q13;q13) and trisomy 8 are nonrandom aberrations in clear-cell sarcoma. Cancer Genet Cytogenet 64: 107-110.

Rohen C, Bonk U, Staats B, Bartnizke S, Bullerdiek J (1993) Two human breast tumors with translocations involving 12q13-15 as the sole cytogenetic abnormality. Cancer Genet Cytogenet 69: 68-71.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sandros J, Stenman G, Mark J (1990): Cytogenetic and molecular observations in human and experimental salivary gland tumours. Cancer Genet Cytogenet 44: 153-167.

Schoenmakers HFFM, Kools PFJ, Kazmierczak B, Bullerdiek J, Claussen U, Horsthemke B, Van den Berghe H, Van de Ven WJM (1993) Isolation of a somatic cell hybrid retaining the der(16)t(12;16)(q13;p11.2) from a myxoid liposarcoma cell line. Cell Genet Cytogenet 62: 159-161.

Schoenmakers HFFM, Kools PFJ, Mols R, Kazmierczak B, Bartnitzke S, Bullerdiek J, Dal Cin P, De Jong PJ, Van den Berghe H, Van de Ven WJM (1994a) Physical mapping of chromosome 12q breakpoints in lipoma, pleomorphic salivary gland adenoma, uterine leiomyoma, and myxoid liposarcoma. Genomics 20: 210-222.

Schoenmakers HFFM, Mols R, Wanschura S, Kools PFJ, Geurts JMW, Bartnitzke S, Bullerdiek J, Van den Berghe H, Van de

Ven WJM (1994b) Identification, molecular cloning and characterization of the chromosome 12 breakpoint cluster region of uterine leiomyomas. *Genes Chromosom Cancer* 11: 106-118.

Second International Chromosome 12 Workshop, New Haven, CT, USA, June 20-22, 1994.

Seifert G, Miehlike A, Haubrich J, Chilla R (1986): Diseases of the salivary glands. Pathology. Diagnosis. Treatment. Facial nerve surgery. Translated by P.M. Stell. Thieme, Stuttgart, New York, pp 182-194.

Smit VTHBM, Wessels JW, Mollevanger P, Schrier PI, Raap AK, Beverstock GC, Cornelisse CJ (1990) Combined GTG-banding and nonradioactive in situ hybridization improves characterization of complex karyotypes. *Cytogenet Cell Genet* 54:20-23.

Sreekantaiah C, Leong SLP, Karakousis CP, McGee DL, Rappaport WD, Villar HV, Neal D, Fleming S, Wankel A, Herrington PN, Carmona R, Sandberg AA (1991) Cytogenetic profile of 109 lipomas. *Cancer Res* 51: 422-433.

Vanni R, Dal Cin P, Marras S, Moerman Ph, Andria M, Valdes E, Deprest J, and Van den Berghe H (1993) Endometrial polyp: Another benign tumor characterized by 12q13-15 changes. *Cancer Genet Cytogenet* 68:32-33.

Walter TA, Xuan Fan S, Medchill MT, Berger CS, Decker H-JH, Sandberg AA (1989) Inv(12)(p11.2q13) in an endometrial polyp. *Cancer Genet Cytogenet* 41: 99-103.

Wanschura S, Belge G, Stenman G, Kools P, Dal Chin P, Schoenmakers E, Huysmans C, Bartnizke S, Van de Ven W, and Bullerdiel J (submitted for publication). Mapping of the translocation breakpoints of primary pleomorphic

adenomas and lipomas within a common region of chromosome  
12.

## LEGENDS OF FIGURES OF ANNEX 1

**Figure 1.** Schematic representation of FISH mapping data obtained for the seven pleomorphic salivary gland adenoma cell lines tested in this study. Cosmid clones which were used as probes in the FISH mapping studies map at sequence-tagged sites obtained from overlapping YAC clones. They are named after the acronyms of the STSs, as shown in the boxes, and the relative order of these is as presented. The DNA interval between RM69 and RM72 is estimated to be about 2.8 Mb. The solid lines indicate DNA intervals in which the breakpoints of the various cell lines are located. The dots indicate FISH experiments that were performed on metaphase chromosomes of the various cell lines using a cosmid clone corresponding to the STS indicated above these as molecular probe. The relative positions of MAR and ULCR12 are indicated in the lower part of the figure. Ad, pleomorphic salivary gland adenoma; MAR, multiple aberration region; ULCR12, uterine leiomyoma cluster region of chromosome 12 breakpoints.

**Figure 2.** a: Partial karyotype of Ad-295/SV40 showing der(8), der(12), der(18) and the corresponding normal chromosomes. b: FISH analysis of metaphase chromosomes of Ad-295/SV40 cells using DNA of cosmid clone cRM76 as molecular probe. Hybridization signals on normal chromosome 12 (arrow) and der(12) (arrowhead). c: GTG-banding pattern of metaphase chromosomes of Ad-295/SV40 shown in b. d: FISH analysis of metaphase chromosomes of Ad-295/SV40 cells using DNA of cosmid clone cRM103 as molecular probe. Hybridization signals on normal chromosome 12 (arrow) and der(18) (arrowhead).

**Figure 3.** Schematic representation of chromosome 12 breakpoint mapping data obtained for primary pleomorphic salivary gland adenomas, uterine leiomyomas, and lipomas as well as cell lines derived from such solid tumors. Results are compared to data for primary uterine leiomyomas (Wanschura et al., submitted for publication)

and cell lines derived from such tumors (Schoenmakers et al., 1994b). Cosmid clones which were used as probes in the FISH mapping studies correspond to sequence-tagged sites obtained from overlapping YAC clones. Cosmid clones 5 were named after the acronyms of the STSs, as shown in the boxes, and the relative order of these is as presented. The estimated sizes of DNA intervals between STSs are indicated. Ad, pleomorphic salivary gland adenoma; Li, lipoma; LM, uterine leiomyoma.

## ANNEX 2

## LEAD ARTICLE

## Identification of the Chromosome 12 Translocation

- 5 Breakpoint Region of a Pleomorphic Salivary Gland Adenoma  
with t(1;12)(p22;q15) as the Sole Cytogenetic Abnormality

Patrick F.J. Kools, Sylke Wanschura, Eric F.P.M.

Schoenmakers, Jan W.M. Geurts, Raf Mols, Bernd

- 10 Kazmierczak, Jörn Bullerdiek, Herman Van den Berghe and  
Wim J.M. Van de Ven

**ABSTRACT:** Cell line Ad-312/SV40, which was derived from a  
primary pleomorphic salivary gland adenoma with  
15 t(1;12)(p22;q15), was used in fluorescence in situ  
hybridization (FISH) analysis to characterize its  
translocation breakpoint region on chromosome 12. Results  
of previous studies have indicated that the chromosome 12  
breakpoint in Ad-312/SV40 is located proximally to locus  
20 D12S8 and distally to the CHOP gene. We here describe two  
partially overlapping yeast artificial chromosome (YAC)  
clones, Y4854 (500 kbp) and Y9091 (460 kbp), which we  
isolated in the context of a chromosome walking project  
with D12S8 and CHOP as starting points. Subsequently, we  
25 have isolated cosmid clones corresponding to various  
sequence-tagged sites (STSs) mapping within the inserts  
of these YAC clones. These included cRM51, cRM69, cRM85,  
cRM90, cRM91, cRM110, and cRM111.

- 30 From the Center for Human Genetics (P.F.J.K., E.F.P.M.S., J.W.M.G., R.M., H.V.D.B.,  
W.J.M.V.D.V.), University of Leuven, Leuven, Belgium and the Center for Human Genetics  
(S.W., B.K., J.B.), University of Bremen, Bremen, Germany.

P.F.J. Kools and Sylke Wanschura contributed equally to this study and must be considered joint  
first authors.

- 35 Address reprint requests to Dr. Wim J.M. Van de Ven, Center for Human Genetics, University of  
Leuven, Herestraat 49, B-3000. Leuven, Belgium.

Received April 13, 1994; accepted July 6, 1994.

We present a composite long-range restriction map encompassing the inserts of these two YAC clones and show by FISH analysis, that both YACs span the chromosome 12 breakpoint as present in Ad-312/SV40 cells. In FISH studies, cosmid clones CRM85, CRM90 and CRM111 appeared to map distally to the chromosome 12 breakpoint whereas cosmid clones CRM51, CRM69, CRM91, and CRM110 were found to map proximally to it. These results assign the chromosome 12 breakpoint in Ad-312/SV40 to a DNA region of less than 165 kbp. FISH evaluation of the chromosome 12 breakpoints in five other pleomorphic salivary gland adenoma cell lines indicated that these are located proximally to the one in Ad-312/SV40, at a distance of more than 0.9 Mb from STS RM91. These results, while pinpointing a potentially critical region on chromosome 12, also provide evidence for the possible involvement of chromosome 12q13-q15 sequences located elsewhere.

#### INTRODUCTION

Pleomorphic salivary gland adenoma constitutes a benign epithelial tumor that originates from the major and minor salivary glands. It is the most common type of salivary gland tumor and accounts for almost 50% of all neoplasms in these organs; 85% of the tumors are found in the parotid gland, 10% in the minor salivary glands, and 5% in the submandibular gland [1]. About 50% of these adenomas appear to have a normal karyotype but cytogenetic studies have also revealed recurrent specific chromosome anomalies [2, 3]. Frequently observed anomalies include aberrations of chromosome 8, usually involving the 8q12-q13 region, with the most common aberration a t(3;8)(p21;q12), and aberrations of chromosome 12, usually translocations involving region 12q13-q15. Non-recurrent clonal chromosome abnormalities have also been reported. The highly specific pattern of chromosome rearrangements with consistent breakpoints at 8q12-q13 and 12q13-q15 suggests that these chromosomal regions harbour genes that might be implicated in the

development of these tumors. Molecular cloning of the chromosome breakpoints and characterization of their junction fragments may lead to the identification of pathogenetically relevant genes. At present, no such  
5 molecular data have yet been reported for these tumors.

On the basis of fluorescence in situ hybridization (FISH) data, the chromosome 12 breakpoints in six pleomorphic salivary gland adenoma cell lines were recently shown to be mapping to region 12q13-q15, more  
10 precisely, to the genomic interval between loci D12S19 and D12S8 [4, 5]. The sex-averaged genetic size of this genomic DNA interval was reported at HGM10 to be 7 cM [6]. We also reported that the chromosome 12 breakpoints in salivary gland adenomas map distally to the CHOP gene  
15 [5], which supports an earlier study indicating that the 12q13-q15 translocation breakpoints in pleomorphic salivary gland adenomas are different from that in myxoid liposarcoma [7]. Here, we report about the physical mapping of the chromosome 12 breakpoint in pleomorphic  
20 salivary gland adenoma cell line Ad-312/SV40, which carries a t(1;12)(p22;q15) as the only cytogenetic abnormality.

#### MATERIALS AND METHODS

##### 25 Tumor Cell Lines.

Human tumor cell lines used in this study included the previously described pleomorphic salivary gland adenoma cell lines Ad-248/SV40, Ad-263/SV40, Ad-295/SV40, Ad-302/SV40, Ad-312/SV40, and Ad-366/SV40 [5,  
30 8]. Cells were cultivated in TC199 culture medium with Earle's salts supplemented with 20% fetal bovine serum. Other cell lines used in this study included somatic cell hybrid PK89-12, which contains chromosome 12 as the sole human chromosome in a hamster genetic background [9], and  
35 somatic cell hybrid LIS-3/SV40/A9-B4 [4]. The latter cell line was obtained upon fusion of myxoid liposarcoma cell line LIS-3/SV40, which carries the specific t(12;16)(q13;p11.2), with mouse A9 cells. This somatic

cell hybrid was previously shown to contain der(16) but neither der(12) nor the normal chromosome 12 [4]. PK89-12 and LIS-3/SV40/A9-B4 cells were grown in DME-F12 medium supplemented with 10% fetal bovine serum. Cell lines were  
5 analyzed by standard cytogenetic techniques at regular intervals.

#### Isolation of YAC and Cosmid Clones.

In the context of human genome mapping studies,  
10 which will be described in detail elsewhere (Schoenmakers et al., in preparation), we isolated YAC clones Y4854 and Y9091 from the first-generation CEPH YAC library [10], and cosmid clones cRM51, cRM69, cRM85, cRM90, cRM91, cRM103, cRM110, and cRM111 from the chromosome-12-  
15 specific, arrayed cosmid library LLNLNC01 [11]. YAC and cosmid clones were isolated as described before [5]. Initial screenings of the YAC, as well as the cosmid library, were performed using a screening strategy involving the polymerase chain reaction (PCR) [12].  
20 Filter hybridization analysis was used as the final screening step, as previously described [5]. Cosmid clones were isolated using STSs and those corresponding to STSs within the inserts of YAC clones Y4854 and Y9091 are indicated in Fig. 1. STSs were obtained via rescue of  
25 YAC insert end-sequences using a vectorette-PCR procedure [13] or Alu-PCR [14, 15]. PCR products were sequenced directly via solid-phase fluorescent sequencing. Cosmid clones were grown and handled according to standard procedures [16]. YAC clones were characterized by pulsed-  
30 field gel electrophoresis [17], restriction mapping, and hybridization, as previously described [5].

#### Chromosome Preparations and Fluorescence in Situ Hybridization.

35 Cells from the pleomorphic salivary gland adenoma tumor cell lines were treated with Colcemid (0.04  $\mu$ g/ml) for 30 min and then harvested according to routine methods. Metaphase spreads of the tumor cells were

prepared as described before [4]. To establish the identity of chromosomes in the FISH experiments, FISH analysis was performed after G-banding of the same metaphase spreads. G-banding was performed essentially as described by Smit et al. [18]. In situ hybridizations were carried out according to a protocol described by Kievits et al. [19] with some minor modifications [5, 20]. Cosmid and YAC DNA was labelled with biotin-11-dUTP (Boehringer Mannheim) or biotin-14-dATP (BRL, Gaithersburg), as described earlier [5]. Chromosomes were counterstained with propidium iodide and analyzed on a Zeiss Axiophot fluorescence microscope using a FITC filter (Zeiss). Results were recorded on Scotch (3M) 640asa film.

## RESULTS

### Isolation and Characterization of YAC Clones Spanning the Chromosome 12 Breakpoint of Pleomorphic Salivary Gland Adenoma Cell Line Ad-312/SV40.

In previous studies [5], we mapped the chromosome 12 breakpoints of six pleomorphic salivary gland adenoma cell lines proximally to locus D12S8 and distally to CHOP. The DNA interval between these loci is somewhat smaller than 7 cM (estimated distance between the loci D12S8 and D12S19 [6]) but still substantially large. To molecularly define the translocation breakpoint of Ad-312/SV40, we have performed human genome mapping studies on the DNA interval between locus D12S8 and the CHOP gene. In the process of directional chromosome walking starting from D12S8 and the CHOP gene, we obtained overlapping YAC clones Y9091 and Y4854. The DNA insert of Y9091 appeared to be 460 kbp and that of Y4854, 500 kbp. Moreover, as we will demonstrate below, the DNA insert of each YAC clone appeared to span the chromosome 12 breakpoint of Ad-312/SV40. A long-range restriction map of the inserts of these YAC clones was made using pulsed-field gel electrophoresis and hybridization analysis (Fig. 1). On the basis of STS content mapping

and Southern blot analysis, the inserts of YAC clones Y9091 and Y4854 appeared to overlap as indicated in Figure 1. The tested STSS correspond to end-sequences of other overlapping YAC clones not shown here or to  
5 sequences obtained via inter-Alu-PCR. Of these, RM90 and RM91 represent such end-clone STSS of YAC Y9091, and RM48 and RM54 of Y4854, whereas RM110 and RM111 represent STSS derived from inter-Alu-PCR. For a number of STSS mapping within the inserts of YAC clones Y4854 and Y9091,  
10 corresponding cosmid clones were isolated for use in FISH analysis, e.g., cRM51, cRM69, cRM85, cRM90, cRM91, cRM110, and cRM111.

The inserts of the two overlapping YAC clones are most likely not chimeric, as was deduced from the  
15 following observations. FISH analysis of metaphase chromosomes of normal human lymphocytes with Y4854 or Y9091 DNA as molecular probe revealed hybridization signals only in chromosome region 12q13-q15. For Y9091, this was confirmed further by observations made in FISH  
20 studies in which cosmid clone cRM90 or cRM91 was used as probe; the DNA insert of each of these two cosmids corresponds to the alternative end-sequences of YAC clone Y9091. Finally, the end-sequence STSS of Y9091 appeared to map to chromosome 12 and distally to the CHOP gene, as  
25 was established by PCR analysis on PK89-12 DNA, which contains human chromosome 12 as the sole human chromosome in a hamster genetic background, and LIS-3/SV40/A9-B4 DNA, which was previously shown to contain der(16), from the specific t(12;16) of myxoid liposarcoma, but neither  
30 der(12) nor the normal chromosome 12 [4]. From the chromosome walking studies, we concluded that the overlapping inserts of the two YAC clones represent a DNA region of about 640-kbp, which is located on chromosome 12q between D12S8 and CHOP. As the 640-kbp composite  
35 long-range restriction map of the YAC contig was constructed with at least double coverage of the entire region, it is not unreasonable to assume that the 640-kbp

region is contiguous with the chromosomal DNA, although microdeletions can not be excluded at this point.

Chromosome walking was routinely evaluated by FISH mapping of YAC clones and/or cosmid clones corresponding to YAC insert sequences. It should be noted that for the identification of chromosomes, G-banding was used in most cases. On the basis of such G-banding, hybridization signals could be assigned conclusively to chromosomes of known identity; this was also of importance for the cases with cross- or background hybridization signals that were occasionally observed. G-banding prior to FISH analysis resulted sometimes in rather weak hybridization signals or rather vague banding patterns. Therefore, we performed FISH experiments in which the YAC and cosmid clones to be evaluated were used in combination with a reference probe. Cosmid clone cPK12qter, which was serendipitously obtained during screening of a cosmid library, was selected as reference marker. FISH analysis of metaphase chromosomes of normal lymphocytes (Fig. 2A) revealed that cPK12qter maps to the telomeric region of the long arm of chromosome 12. To identify chromosome 12 in this experiment, centromere 12-specific probe p $\alpha$ 12H8 [21] was used. FISH analysis of metaphase chromosomes of Ad-312/SV40 cells using YAC clone Y4854 (Fig. 2B) or Y9091 (Fig. 2C) in combination with reference probe cPK12qter revealed, in both cases, hybridization signals of the YAC insert on der(1) as well as der(12). We concluded from these results that the insert DNA of each YAC clone might span the chromosome 12 breakpoint in this cell line. It should be noted that G-banding revealed a telomeric association involving the short arm of chromosome 12 in Figure 2C. The observation that YAC clone Y9091 spanned the chromosome 12 breakpoint in Ad-312/SV40 was confirmed independently in FISH studies in which cosmid clone cRM90 or cRM91 was used as molecular probe; they were shown to contain the alternative end-sequences of the Y9091 insert. cRM90 appeared to map distally to the chromosome 12 breakpoint,

whereas CRM91 was found to map proximally (data not shown). These results also established the chromosomal orientation of the YAC contig shown in Figure 1. In summary, we concluded from these FISH studies that the chromosome 12 translocation breakpoint in Ad-312/SV40 must be located in the DNA interval corresponding to the overlapping sequences (about 300 kbp) of the two YAC clones.

10 **Fine Mapping of the Chromosome 12 Translocation  
Breakpoint of Ad-312/SV40.**

In an approach to further narrow the chromosome 12 translocation breakpoint region of Ad-312/SV40, cosmid clones with different mapping positions within YAC clone Y9091 were isolated. These included CRM69, CRM85, CRM110, and CRM111. CRM69 and CRM85 were isolated on the basis of STS sequences of YAC clones not shown here. CRM110 and CRM111 were obtained via inter-Alu-PCR. RM110 was shown by Southern blot analysis to hybridize to a terminal MluI fragment of Y9091 and not to the DNA insert of the overlapping YAC clone with RM69 as telomeric end-sequences. The location of RM110 is as indicated in Figure 1. RM111 was shown to hybridize to a BssHII, MluI, PvuI, and SfiI fragment of Y9091 and is therefore located in the PvuI-SfiI fragment of Y9091, to which STS RM48 was also mapped (Fig. 1). FISH analysis of metaphase chromosomes of Ad-312/SV40 with CRM69 or CRM110 as probe indicated that the DNA insert of these cosmids mapped proximally to the chromosome 12 translocation breakpoint in this cell line, as illustrated for CRM69 in Figure 3A. Subsequent FISH analysis of Ad-312/SV40 with CRM85 or CRM111 as probe revealed hybridization signals distally to the translocation breakpoint, as illustrated for CRM111 in Figure 3B. The results with CRM85 and CRM111 are in agreement with the observed breakpoint spanning by YAC clone Y4854 as CRM85 maps distally and CRM111 closely to STS RM48, which marks the telomeric end of YAC clone Y4854. In conclusion, the chromosome 12 translocation

breakpoint in Ad-312/SV40 must be located in the DNA interval between cRM110 and cRM111, as schematically summarized in Figure 4.

5 **FISH Evaluation of Chromosome 12 Breakpoints in Other Pleomorphic Salivary Gland Adenoma Cell Lines.**

To determine the position of their chromosome 12 breakpoints relative to that of Ad-312/SV40, five other pleomorphic salivary gland adenoma cell lines were  
10 evaluated by FISH analysis, as summarized schematically in Figure 4. These cell lines, which were developed from primary tumors [5, 8], included Ad-248/SV40, Ad-263/SV40, Ad-295/SV40, Ad-302/SV40, and Ad-366/SV40. The chromosome 12 aberrations of these cell lines are listed in Figure  
15 4. FISH analysis of metaphase chromosomes of these cell lines using cRM91 revealed that the chromosome 12 breakpoints of all these cell lines mapped proximally to this cosmid clone (data not shown). Similar FISH analysis was also performed using a cosmid clone corresponding to  
20 sequence-tagged site RM103 as a probe. RM103 was found to map proximally to RM91 at a distance of about 0.9 Mbp. In all cases, cRM103 appeared to map distally to the chromosome 12 translocation breakpoints, indicating that the chromosome 12 breakpoints in these five pleomorphic  
25 salivary gland adenoma cell lines are located at a relatively large distance from that of Ad-312/SV40 cells.

**DISCUSSION**

In the studies presented here, we have  
30 identified, molecularly cloned, and characterized a chromosome region on the long arm of chromosome 12 in which the translocation breakpoint of pleomorphic salivary gland adenoma cell line Ad-312/SV40 appears to map. In previous studies [5], we already provided  
35 evidence that the chromosome 12 breakpoint of this cell line was located between D12S8 and CHOP. Because the two breakpoints spanning YAC clones described here were obtained in directional chromosome walking experiments

using D12S8 and the CHOP gene as initial starting points, the chromosome 12 breakpoint mapping results presented here confirm our previous claim. The FISH results obtained with the complete YAC insert of Y9091 as  
5 molecular probe were confirmed independently in FISH studies using cosmid clones containing sequences corresponding to various regions of the insert of this YAC clone. This is of importance, as the independent confirmatory results make it rather unlikely that the  
10 split signals observed with the complete insert of Y9091 can be explained otherwise than by a factual splitting of sequences represented in the YAC. The presence, for instance, of highly related genetic sequences on both sides of a chromosome breakpoint could easily lead to  
15 erroneous conclusions if they were based solely on FISH results of a YAC insert. Finally, our mapping studies have also established conclusively the chromosomal orientation of the long-range restriction map we have generated in these studies. This orientation was already  
20 predicted on the basis of two-color FISH studies (unpublished observations).

The FISH studies, described here, enabled us to map the chromosome 12 breakpoint in Ad-312/SV40 cells to the 190-kbp DNA interval between the established STS  
25 RM48 and RM69. However, the breakpoint region can be narrowed somewhat further on the basis of the following. The fact that Y4854 was shown to span the breakpoint indicates that at least a considerable part of the telomeric half of this YAC clone must map distally to the  
30 breakpoint. Precisely how much remains to be established. On the other side, STS RM69 appeared to be located in about the middle of the DNA insert of cosmid clone cRM69, suggesting that the breakpoint is close to 25 kbp distally to RM69. Moreover, cRM69 appeared to lack RM110  
35 (data not shown) and, as cRM110 was found proximally to the chromosome 12 breakpoint in Ad-312/SV40 cells, the breakpoint should be even further distal to RM69 than the earlier-mentioned 25 kbp. Altogether, this narrows the

chromosome 12 breakpoint region to a DNA interval, which must be considerably smaller than 165 kbp. Further pinpointing of the breakpoint will allow us to molecularly clone the chromosome 12 breakpoint and to  
5 characterize the genetic sequences in the breakpoint junction region, which might lead to the identification of pathogenetically relevant sequences. Identification of the genes present in the DNA inserts of YAC clones Y4854 and Y9091, via sequencing, direct hybridization, direct  
10 selection or exon-trapping, might constitute a useful alternative approach for identifying the gene in this region of the long arm of chromosome 12 that might be pathogenetically critical for pleomorphic salivary gland adenoma tumorigenesis.

15 The observation that the chromosome 12 breakpoints in other pleomorphic salivary gland adenomas are located in a remote and more proximal region on the long arm of chromosome 12 is of interest. It could imply that the chromosome 12 breakpoints in pleomorphic  
20 salivary gland adenomas are dispersed over a relatively large DNA region of the long arm of chromosome 12, reminiscent to the 11q13 breakpoints in B-cell malignancies [22]. Elucidation of the precise location of the chromosome 12 breakpoints in the other pleomorphic  
25 salivary gland adenoma cell lines could shed more light on this matter. On the other hand, it could point towards alternative sequences on the long arm of chromosome 12 between D12S8 and the CHOP gene that might be of importance, presumably for growth regulation in  
30 pleomorphic salivary gland adenoma. The fact that the chromosome 12 breakpoint region described here has so far been found only in the Ad-312/SV40 cell line makes it necessary to analyze a larger number of salivary gland adenomas with chromosome 12q13-q15 aberrations to assess  
35 the potential relevance for tumorigenesis of the chromosome 12 sequences affected in the studied cell line. If more cases with aberrations in this particular region of chromosome 12 can be found, it would be of

interest to find out whether these tumors form a clinical subgroup. Finally, chromosome translocations involving region q13-q15 of human chromosome 12 have been reported for a variety of other solid tumors: benign adipose  
5 tissue tumors, uterine leiomyoma, rhabdomyosarcoma, hemangiopericytoma, clear-cell sarcoma, chondromatous tumors, and hamartoma of the lung. Whether or not the chromosome 12 breakpoints in some of these tumors map within the same region as that of Ad-312/SV40 remains to  
10 be established. The YAC and cosmid clones described in this report constitute useful tools to investigate this.

The availability of a copy of the first-generation CEPH YAC library [10] and a copy of the  
15 arrayed chromosome 12-specific cosmid library (LLNL12NC01) [11] is greatly acknowledged. The cosmid library was constructed as part of the National Laboratory Gene Library Project under the auspices of the U.S. DOE by LLNL under contract No. W-7405-Eng-48. The  
20 authors acknowledge the excellent technical assistance of M. Dehaen, C. Huysmans, E. Meyen, K. Meyer-Bolte, and M. Willems and would like to thank M. Leys for art work. This work was supported in part by the EC through Biomed  
1 program "Molecular Cytogenetics of Solid Tumours", the  
25 "Geconcerteerde Onderzoekacties 1992-1996", the "Association Luxembourgeoise contre le Cancer", the National Fund for Scientific Research (NFWO; Kom op tegen Kanker), the "ASLK-programma voor Kankeronderzoek", the "Schwerpunktprogramm: Molekulare und Klassische  
30 Tumorcytogenetik" of the Deutsche Forschungsgemeinschaft, and the Tönjes-Vagt Stiftung. This text presents results of the Belgian programme on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The  
35 scientific responsibility is assumed by its authors. J.W.M. Geurts is an "Aspirant" of the National Fund for Scientific Research (NFWO; Kom op tegen Kanker).

## REFERENCES

1. Seifert G, Miehlik A, Haubrich J, Chilla R (1986):  
Diseases of the salivary glands. Pathology.  
5 Diagnosis. Treatment. Facial nerve surgery.  
Translated by P.M. Stell. Thieme, Stuttgart, New  
York, pp 1820194.
2. Sandros J, Stenman G, Mark J (1990): Cytogenetic and  
10 molecular observations in human and experimental  
salivary gland tumours. Cancer Genet Cytogenet 44:  
153-167.
3. Bullerdiek J, Wobst G, Meyer-Bolte K, Chilla R,  
15 Haubrich J, Thode B, Bartnitzke S (1993):  
Cytogenetic subtyping of 220 salivary gland  
pleomorphic adenomas: correlation to occurrence,  
histological subtype, and in vitro cellular  
behavior. Cancer Genet Cytogenet 65: 27-31.  
20
4. Schoenmakers HFFM, Kools PFJ, Kazmierczak B,  
Bullerdiek J, Claussen U, Horsthemke B, Van den  
Berghe H, Van de Ven WJM (1993): Isolation of a  
somatic cell hybrid retaining the  
25 der(16)t(12;16)(q13;p11.2) from a myxoid liposarcoma  
cell line. Cell Genet Cytogenet 62: 159-161.
5. Schoenmakers HFFM, Kools PFJ, Mols R, Kazmierczak B,  
Bartnitzke S, Bullerdiek J, Dal Cin P, De Jong PJ,  
30 Van den Berghe H, Van de Ven WJM (1993): Physical  
mapping of chromosome 12q breakpoints in lipoma,  
pleomorphic salivary gland adenoma, uterine  
leiomyoma, and myxoid liposarcoma. Genomics, 20:  
210-222.  
35
6. Keats B, Ott J, Conneally M (1989): Reports of the  
committee on linkage and gene order. Cytogenet Cell  
Genet 51: 459-502.

7. Stenman G, Sahlin P, Mark J, Chaganti RKS, Kindblom LS, Aman P (1993): The 12q13-q15 translocation breakpoints in pleomorphic adenoma and clear-cell sarcoma of tendons and aponeuroses are different from that in myxoid liposarcoma. Genes Chrom Cancer 7: 178-180.
8. Kazmierczak B, Bartnitzke S, Hartl M, Bullerdiek J (1990): In vitro transformation by the SV40 "early region" of cells from a human benign salivary gland tumour with a 12q13-q15 rearrangement. Cytogenet Cell Genet 53: 37-39.
9. Warburton D, Gersen S, Yu M-T, Jackson C, Handelin B, Housman D (1990): Monochromosomal rodent-human hybrids from microcell fusion of human lymphoblastoid cells containing an inserted dominant selectable marker. Genomics 6: 358-366.
10. Albertsen HM, Abderrahim H, Cann HM, Dausset J, Le Paslier D, Cohen D (1990): Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. Proc Natl Acad Sci USA 87: 4256-4260.
11. Montgomery KT, LeBlanc JM, Tsai P, McNinch JS, Ward DC, De Jong PJ, Kucherlapati R, Krauter KS (1993): Characterization of two chromosome 12 cosmid libraries and development of STSs from cosmids mapped by FISH. Genomics 17:682-693.
12. Green ED, Olson MV (1990): Systematic screening of yeast artificial-chromosome libraries using the polymerase chain reaction. Proc Natl Acad Sci USA 87: 1213-1217.

13. Geurts JMW, Schoenmakers HFPM, Mols R, Van de Ven WJM (1994): Improved procedure for rapid isolation and sequencing of DNAs termini in yeast artificial chromosomes. Meth Mol Cell Biol, In Press.
- 5 14. Nelson DL, Ledbetter SA, Corbo L, Victoria MF, Ramirez-Solis R, Webster TD, Ledbetter DH, Caskey CT (1989): Alu polymerase chain reaction. A method for rapid isolation of human-specific sequences from  
10 complex DNA sources. Proc Natl Acad Sci USA 86: 6686-6690.
- 15 15. Breukel C, Wijnen J, Trops C, Van de Klift H, Dauwerse H, Meera Khan P (1990): Vector-Alu PCR: a rapid step in mapping cosmids and YACs. Nucl Acids Res 18: 3097.
16. Sambrook J, Fritsch EF, Maniatis T (1989): Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.  
20
17. Chu G, Vollrath D, Davis RW (1986): Separation of large DNA molecules by contour-clamped homogeneous electric fields. Science 234: 1582-1585.
- 25 18. Smit VTHBM, Wessels JW, Mollevanger P, Schrier PI, Raap EK, Beverstock GC, Cornelisse CJ (1990): Combined GTG-banding and nonradioactive in situ hybridization improves characterization of complex  
30 karyotypes. Cytogenet Cell Genet 54: 20-23.
19. Kievits T, Dauwerse JG, Wiegant J, Devilee P, Breuning MH, Cornelisse CJ, van Ommen G, Pearson PL (1990): Rapid subchromosomal localization of cosmids  
35 by nonradioactive in situ hybridization. Cytogenet Cell Genet 53: 134-136.

20. Kools PFJ, Roebroek AJM, Van de Velde HJK, Marynen  
P, Bullerdiel J, Van de Ven WJM (1993): Regional  
mapping of the human NSP gene to chromosome 14q21-  
q22 by fluorescence in situ hybridization. Cytogenet  
5 Cell Genet 66:48-50.
21. Looijenga LHJ, Smit VTHBM, Wessels JW, Mollevanger  
P, Oosterhuis JW, Cornelisse CJ (1990): Localization  
and polymorphism of a chromosome 12-specific  $\alpha$   
10 satellite DNA sequence. Cytogenet Cell Genet 53:  
216-218.
22. Raynaud SD, Bekri S, Leroux D, Grosgeorge J, Klein  
B, Bastard C, Gaudray P, Simon MP (1993): Expanded  
15 range of 11q13 breakpoints with differing patterns  
of cyclin D1 expression in B-cell malignancies.  
Genes Chrom Cancer 8: 80-87.

## LEGENDS OF FIGURES OF ANNEX 2

**Figure 1.** Composite physical map of the overlapping DNA inserts of YAC clones Y4854 and Y9091. Sizes of the DNA inserts are indicated. The relative positions of the YAC clones are represented by bars below the long range physical map. Sequence-tagged sites (STSs) corresponding to end-clones of YACs, including YACs not shown here, are indicated by boxed RM codes above the restriction map. STSs obtained from inter-Alu-PCR products are given below the restriction map and the DNA regions to which they have been mapped are marked by arrows. B: BssHII; M: MluI; P: PvuI; Sf: SfiI. A polymorphic MluI site is marked by an asterisk.

**Figure 2.** A) Mapping of cosmid clone cPK12qter to the telomeric region of the long arm of chromosome 12. Centromere 12-specific probe p<sub>cl</sub>2H8 was used to establish the identity of chromosome 12. FISH analysis was performed on metaphase chromosomes of control human lymphocytes. Hybridization signals of cPK12qter are marked with small arrowheads, those of the centromere 12-specific probe with asterisks. B, C) FISH analysis of metaphase chromosomes of Ad-312/SV40 cells using DNA of YAC clone Y4854 (B) or Y9091 (C) as molecular probe in combination with cosmid clone cPK12qter as reference marker. Hybridization signals of the YAC clones on chromosome 12 are indicated by large arrowheads; those on der(1) by large arrows, and those on der(12) by small arrows, respectively. The hybridization signals of cosmid clone cPK12qter are indicated by small arrowheads.

**Figure 3.** FISH analysis of metaphase chromosomes of Ad-312/SV40 cells using DNA of cosmid clone cRM69 (A) or cRM111 (B) as molecular probe in combination with cosmid clone cPK12qter as reference marker. The position of the hybridization signals of cPK12qter are indicated by small arrowheads. In (A), the position of the hybridization signal of cRM69 on normal chromosome 12 is indicated by a large arrowhead, and that on der(12) with a small arrow. In (B), the position of the hybridization signal of

cRM111 on normal chromosome 12 is indicated by a large arrowhead, and that on der(1) with a large arrow.

Figure 4. Schematic representation of FISH mapping data obtained for the six pleomorphic salivary gland adenoma cell lines tested in this study. The specific chromosome 12 aberrations in the various cell lines are given. Cosmid clones which were used as probes in the FISH mapping studies correspond to sequence-tagged sites obtained from overlapping YAC clones. Individual FISH experiments are indicated by dots. Cosmid clones were named after the acronyms of the STSs, as shown in the boxes, and the relative order of these is as presented. The DNA interval between RM90 and RM103 is estimated to be about 1.3 Mb. Insert: Schematic representation of the G-banded derivative chromosomes der(1) and der(12) of the Ad-312/SV40 cell line, which carries a t(1;12)(p22;q15). The positions of the chromosome 12 breakpoints of Ad-248/SV40, Ad-263/SV40, Ad-295/SV40, Ad-302/SV40, and Ad-366/SV40 are distal to RM103 as indicated by the arrow.

## CLAIMS

1. Multi-tumor Aberrant Growth (MAG) gene having the nucleotide sequence of any one of the strands of any one of the members of the High Mobility Group protein genes or LIM protein genes, including modified versions thereof.
- 5           2. Multi-tumor Aberrant Growth (MAG) gene as claimed in claim 1 having essentially the nucleotide sequence of the HMGI-C gene as depicted in figure 7, or the complementary strand thereof, including modified or elongated versions of both strands.
- 10           3. Multi-tumor Aberrant Growth (MAG) gene as claimed in claim 1 having essentially the nucleotide sequence of the LPP gene as depicted in figure 5, or the complementary strand thereof, including modified or elongated versions of both strands.
- 15           4. Multi-tumor Aberrant Growth (MAG) gene as claimed in claim 1, 2 or 3 for use as a starting point for designing suitable expression-modulating compounds or techniques for the treatment of non-physiological proliferation phenomena in human or animal.
- 20           5. Multi-tumor Aberrant Growth (MAG) gene as claimed in claim 1, 2 or 3 for use as a starting point for designing suitable nucleotide probes for (clinically/medically) diagnosing cells having a non-physiological proliferative capacity as compared to wildtype  
25 cells.
6. Protein encoded by the Multi-tumor Aberrant Growth (MAG) gene as claimed in claim 1, 2 or 3 for use as a starting point for preparing suitable antibodies for (clinically/medically) diagnosing cells having a non-  
30 physiological proliferative capacity as compared to wildtype cells.
7. Derivatives of the MAG gene as claimed in claim 1, 2 or 3 or of its immediate vicinity for use in diagnosis and the preparation of therapeutical compositions, wherein  
35 the derivatives are selected from the group consisting of sense and anti-sense cDNA or fragments thereof, transcripts

of the gene or practically usable fragments thereof, antisense RNA, fragments of the gene or its complementary strand, proteins encoded by the gene or fragments thereof, antibodies directed to the gene, the cDNA, the transcript, 5 the protein or the fragments thereof, as well as antibody fragments.

8. In situ diagnostic method for diagnosing cells having a non-physiological proliferative capacity, comprising at least some of the following steps:

10 a) designing a set of nucleotide probes based on the information obtainable from the nucleotide sequence of the MAG gene as claimed in claim 1 or 2, wherein one of the probes is hybridisable to a region of the aberrant gene substantially mapping at the same locus as a corresponding 15 region of the wildtype gene and the other probe is hybridisable to a region of the aberrant gene mapping at a different locus than a corresponding region of the wildtype gene;

20 b) incubating one or more interphase or metaphase chromosomes or cells having a non-physiological proliferative capacity, with the probe under hybridising conditions; and

c) visualising the hybridisation between the probe and the gene.

25 9. Method of diagnosing cells having a non-physiological proliferative capacity, comprising at least some of the following steps:

a) taking a biopsy of the cells to be diagnosed;  
b) isolating a suitable MAG gene-related 30 macromolecule therefrom;

c) analysing the macromolecule thus obtained by comparison with a wildtype reference molecule preferably from the same individual.

35 10. Method as claimed in claim 9, comprising at least some of the following steps:

a) taking a biopsy of the cells to be diagnosed;  
b) extracting total RNA thereof;

c) preparing at least one first strand cDNA of the mRNA species in the total RNA extract, which cDNA comprises a suitable tail;

d) performing a PCR and/or RT-PCR using a MAG gene specific primer and a tail-specific and/or partner-specific/nested primer in order to amplify MAG gene specific cDNA's;

e) separating the PCR products on a gel to obtain a pattern of bands;

10 f) evaluating the presence of aberrant bands by comparison to wildtype bands, preferably originating from the same individual.

11. Method as claimed in claim 9, comprising at least some of the following steps:

15 a) taking a biopsy of the cells to be diagnosed;

b) isolating total protein therefrom;

c) separating the total protein on a gel to obtain essentially individual bands and optionally transferring the bands to a Western blot;

20 d) hybridising the bands thus obtained with antibodies directed against a part of the protein encoded by the remaining part of the MAG gene and against a part of the protein encoded by the substitution part of the MAG gene;

e) visualising the antigen-antibody reactions and  
25 establishing the presence of aberrant bands by comparison with bands from wildtype proteins, preferably originating from the same individual.

12. Method as claimed in claim 9, comprising at least some of the following steps:

30 a) taking a biopsy of the cells to be diagnosed;

b) isolating total DNA therefrom;

c) digesting the DNA with one or more so-called "rare cutter" restriction enzymes;

d) separating the digest thus prepared on a gel to  
35 obtain a separation pattern;

e) optionally transferring the separation pattern to a Southern blot;

f) hybridising the separation pattern in the gel or on the blot with one or more informative probes under hybridising conditions;

g) visualising the hybridisations and establishing  
5 the presence of aberrant bands by comparison to wildtype bands, preferably originating from the same individual.

13. Method as claimed in claim 9, comprising at least some of the following steps:

- 10 a) taking a biopsy of the cells to be diagnosed;  
b) extracting mRNA therefrom;  
c) establishing the presence or the (relative) quantity of mRNA derived from the MAG gene; and  
d) comparing the result of step c) with the result of a similar experiment with wildtype cells, preferably  
15 originating from the same individual.

14. Method as claimed in any one of the claims 8-13, wherein the cells having a non-physiological proliferative capacity are selected from the group consisting of the mesenchymal tumors hamartomas (e.g. breast  
20 and lung), adipose tissue tumors (e.g. lipomas), pleomorphic salivary gland adenomas, uterine leiomyomas, angiomyxomas, fibroadenomas of the breast, polyps of the endometrium, atherosclerotic plaques, and other benign tumors as well as various malignant tumors, including but not limited to  
25 sarcomas (e.g. rhabdomyosarcoma, osteosarcoma) and carcinomas (e.g. of breast, lung, skin, thyroid), and haematological malignancies, like leukemias and lymphomas.

15. Anti-sense molecules of a MAG gene as claimed in claim 1, 2 or 3 for use in the treatment of diseases  
30 involving cells having a non-physiological proliferative capacity by modulating the expression of the gene.

16. Expression modulators, such as inhibitors or enhancers, including ribozymes, of the MAG gene as claimed in claim 1, 2 or 3 for use in the treatment of diseases  
35 involving cells having a non-physiological proliferative capacity.

17. Antisense RNA molecules complementary to the mRNA molecules of the MAG gene and/or antibodies directed

against the gene product of the MAG gene as claimed in claim 1, 2 or 3 for use in the treatment of diseases involving cells having a non-physiological proliferative capacity.

18. Diagnostic kit for performing the method as  
5 claimed in claim 8, comprising a suitable set of labeled nucleotide probes.

19. Diagnostic kit for performing the method as claimed in claim 10, comprising a suitable set of labeled probes.

10 20. Diagnostic kit for performing the method as claimed in claim 11, comprising a suitable set of labeled MAG gene specific and tail specific PCR primers.

21. Diagnostic kit for performing the method as claimed in claim 11, comprising a suitable set of labeled  
15 probes, and suitable rare cutting restriction enzymes.

22. Pharmaceutical composition for lowering the expression level of the MAG gene in cells having a non-physiological proliferative capacity, comprising one or more of the derivatives as claimed in claim 7 and/or one or more  
20 expression modulators as claimed in claim 16.

23. Pharmaceutical composition as claimed in claim 22, wherein the cells having a non-physiological proliferative capacity are selected from the group consisting of the mesenchymal tumors hamartomas (e.g. breast  
25 and lung), adipose tissue tumors (e.g. lipomas), pleomorphic salivary gland adenomas, uterine leiomyomas, angiomyxomas, fibroadenomas of the breast, polyps of the endometrium, atherosclerotic plaques, and other benign tumors as well as various malignant tumors, including but not limited to  
30 sarcomas (e.g. rhabdomyosarcoma, osteosarcoma) and carcinomas (e.g. of breast, lung, skin, thyroid), and haematological malignancies, like leukemias and lymphomas.

24. Use of the derivatives as claimed in claim 7 for the preparation of a diagnostic kit or a pharmaceutical  
35 composition for the diagnosis or treatment of diseases or disorders involving cells having a non-physiological proliferative capacity.

25. Use of the expression modulators as claimed in claim 16 for the preparation of a pharmaceutical composition for the treatment of diseases or disorders involving cells having a non-physiological proliferative capacity.

- 5           26. Use as claimed in claim 24 or 25, wherein the cells having a non-physiological proliferative capacity are selected from the group consisting of the mesenchymal tumors hamartomas (e.g. breast and lung), adipose tissue tumors (e.g. lipomas), pleomorphic salivary gland adenomas, uterine  
10 leiomyomas, angiomyxomas, fibroadenomas of the breast, polyps of the endometrium, atherosclerotic plaques, and other benign tumors as well as various malignant tumors, including but not limited to sarcomas (e.g. rhabdomyosarcoma, osteosarcoma) and carcinomas (e.g. of  
15 breast, lung, skin, thyroid), and haematological malignancies, like leukemias and lymphomas.

27. Method for isolating other MAG genes based on the existence of a fusion gene, fusion transcript or fusion protein in a tumor cell by using at least a part of a MAG  
20 gene for designing molecular tools (probes, primers etc.).

28. MAG genes obtainable by the method of claim 27.

29. MAG genes as claimed in claim 28 for use in diagnostic or therapeutic methods.

- 25           30. Animal model for the assessment of the utility of compounds or compositions in the treatment of diseases or disorders involving cells having a non-physiological proliferative capacity, which animal is a transgenic animal harbouring a MAG gene in its genome.

- 30           31. Animal model as claimed in claim 30, wherein the MAG gene is an aberrant MAG gene, such as a fusion product of the remaining part of the gene and the substitution part of its translocation partner.

32. Animal model as claimed in claim 30, wherein  
35 the MAG gene shows a non-physiological expression level.

33. Animal model for the assessment of the utility of compounds or compositions in the treatment of diseases or disorders involving cells having a non-physiological

proliferative capacity, which animal harbours a specific genetic aberration affecting a MAG gene as claimed in claim 1, 2 or 3 in the genome of at least part of its cells, which aberration is induced via homologous recombination in 5 embryonic stem cells.

34. Animal model as claimed in any one of the claims 30-33, which animal is a mammal, in particular a mouse, rat, dog, pig or higher primate, like chimpanzee.

35. Poly- or oligonucleotide probes and primers as 10 disclosed in the description and figures.

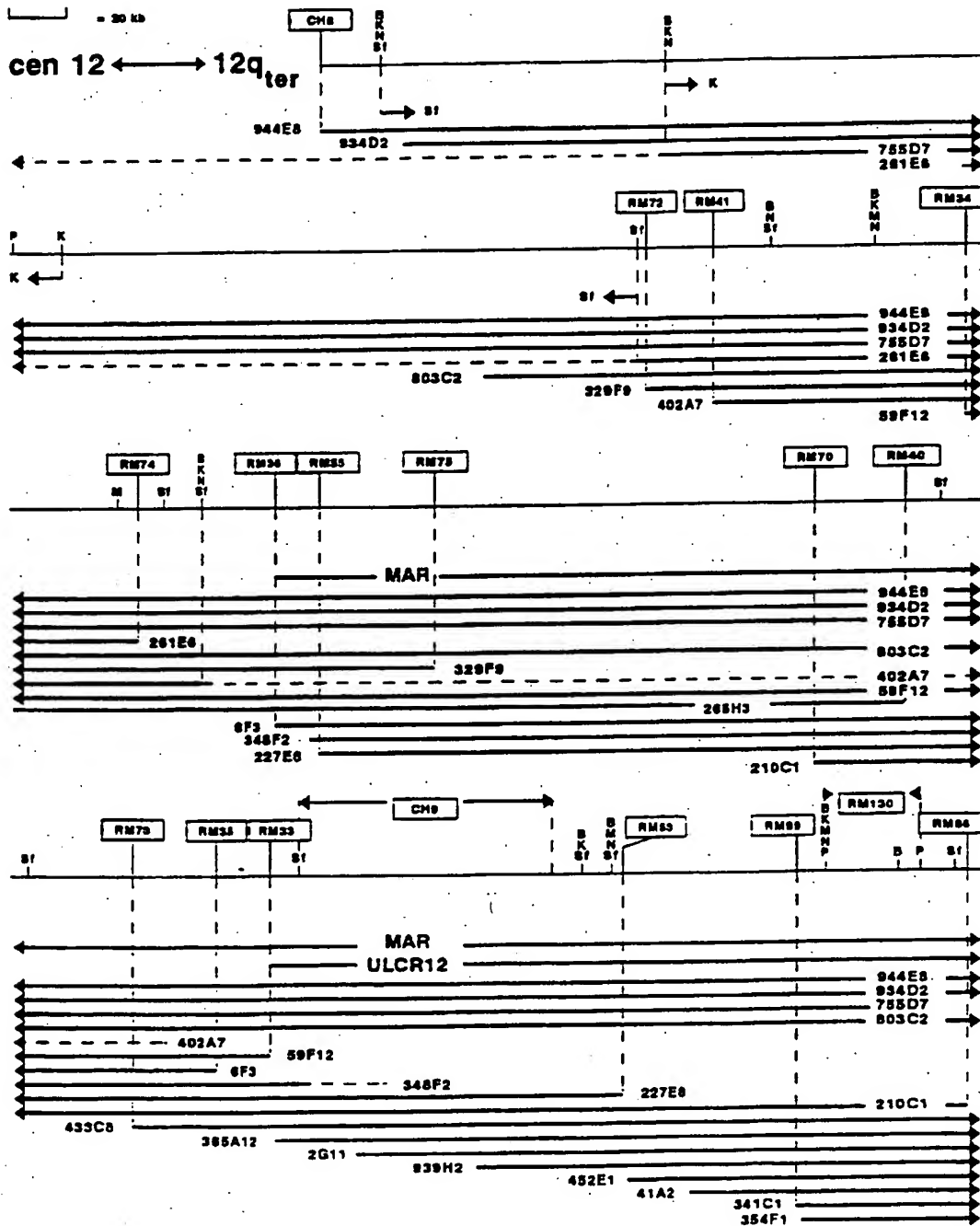


FIG. 1-1

2/21

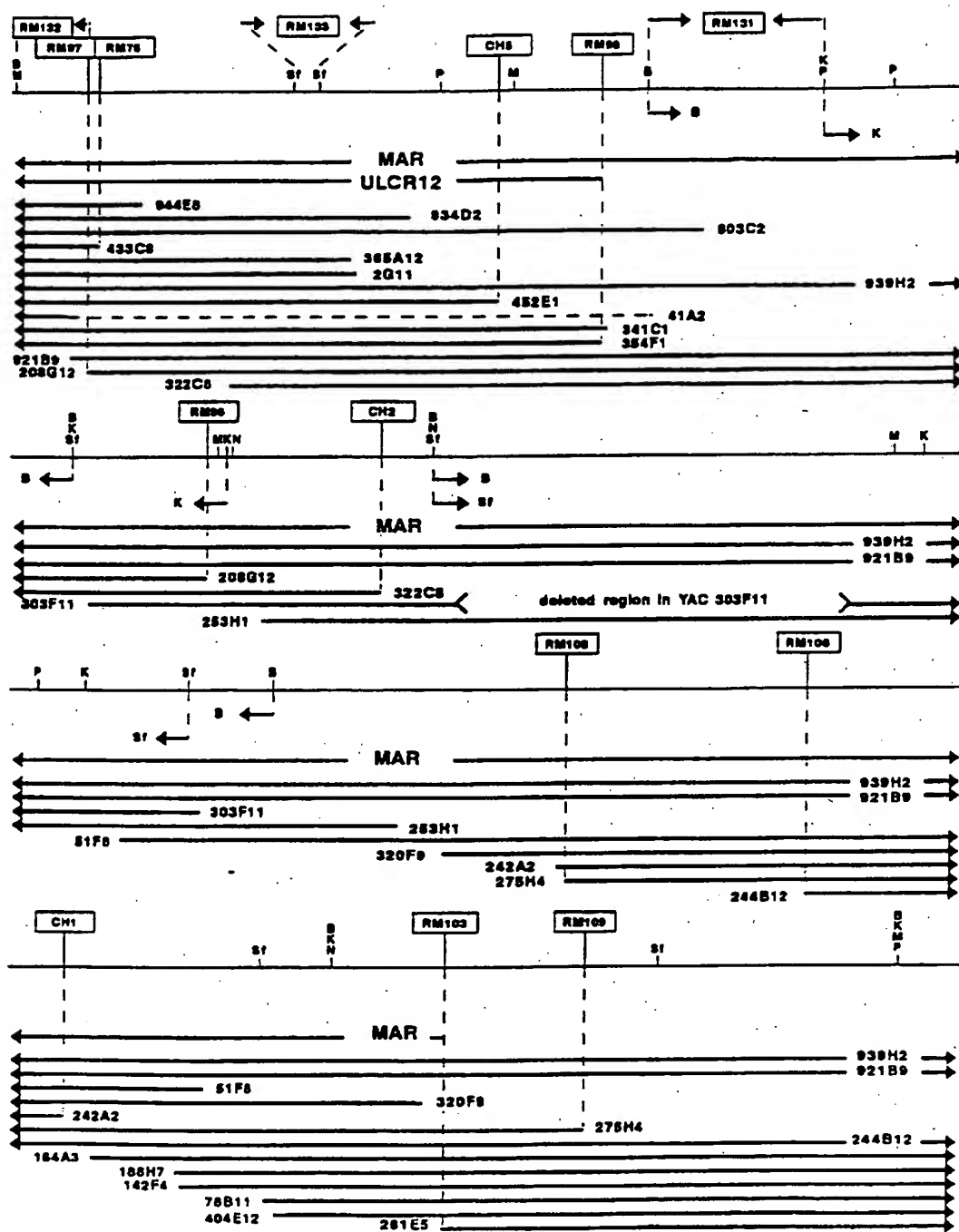


FIG. 1-2

3/21

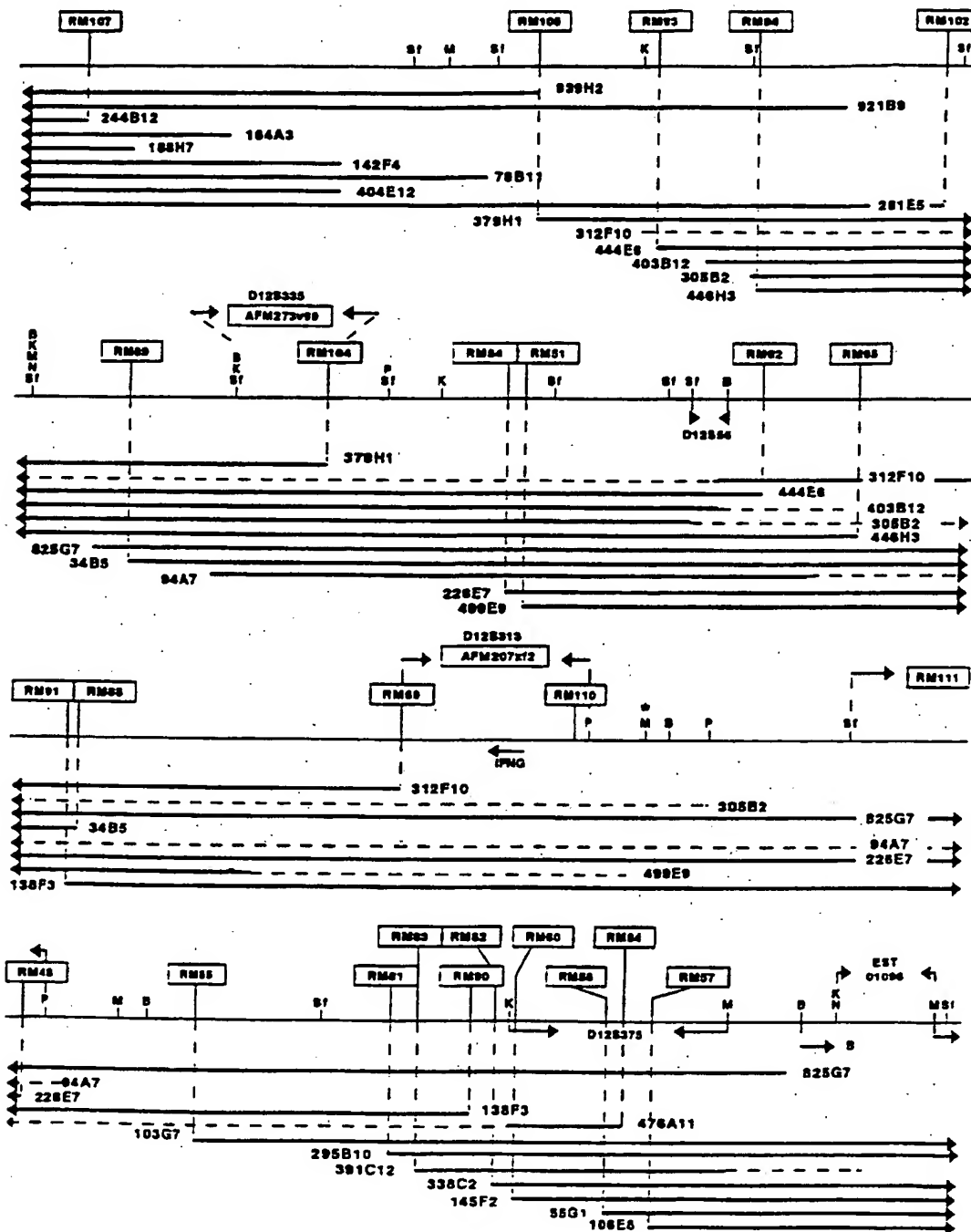


FIG. 1-3

4/21

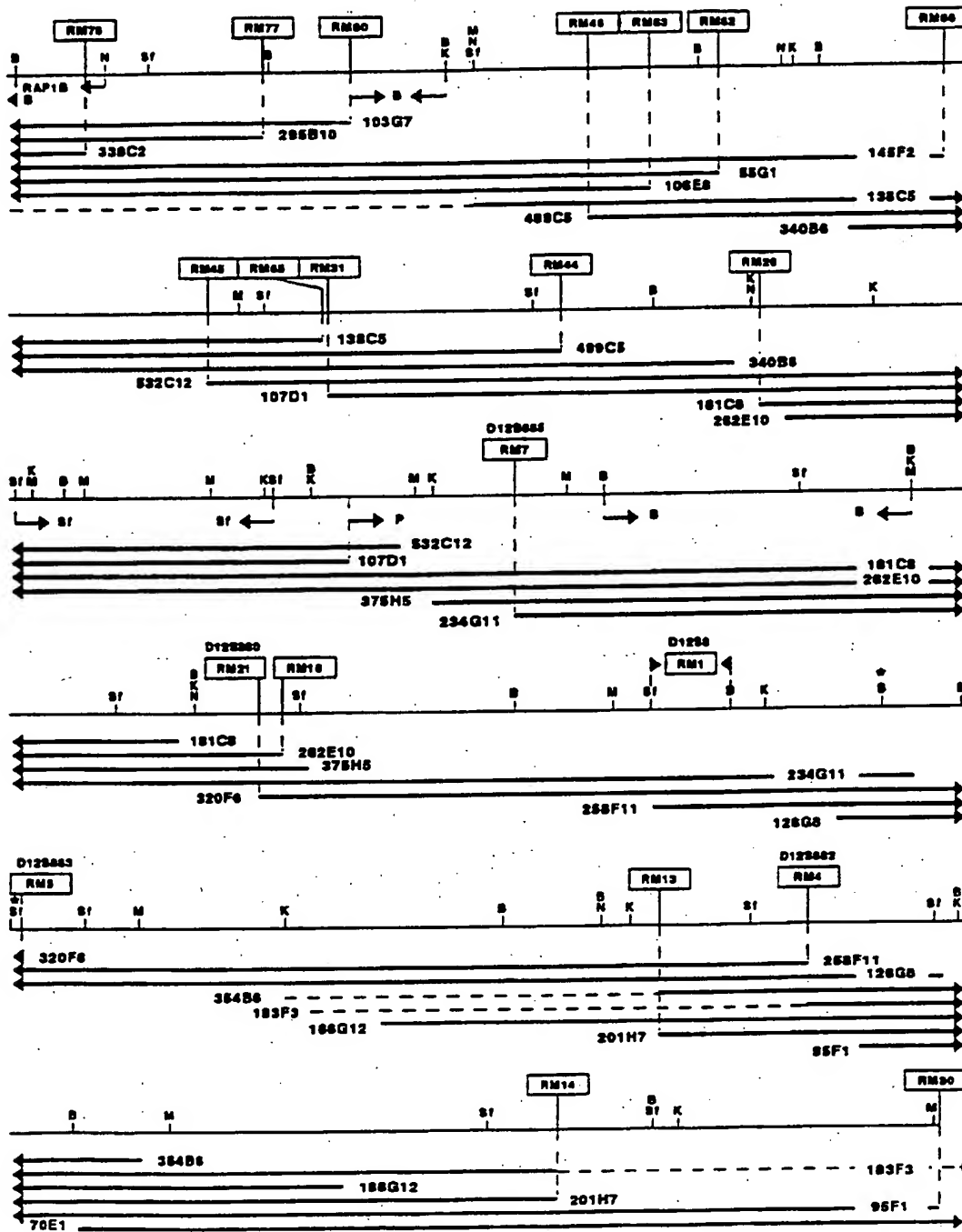


FIG. 1-4

5/21

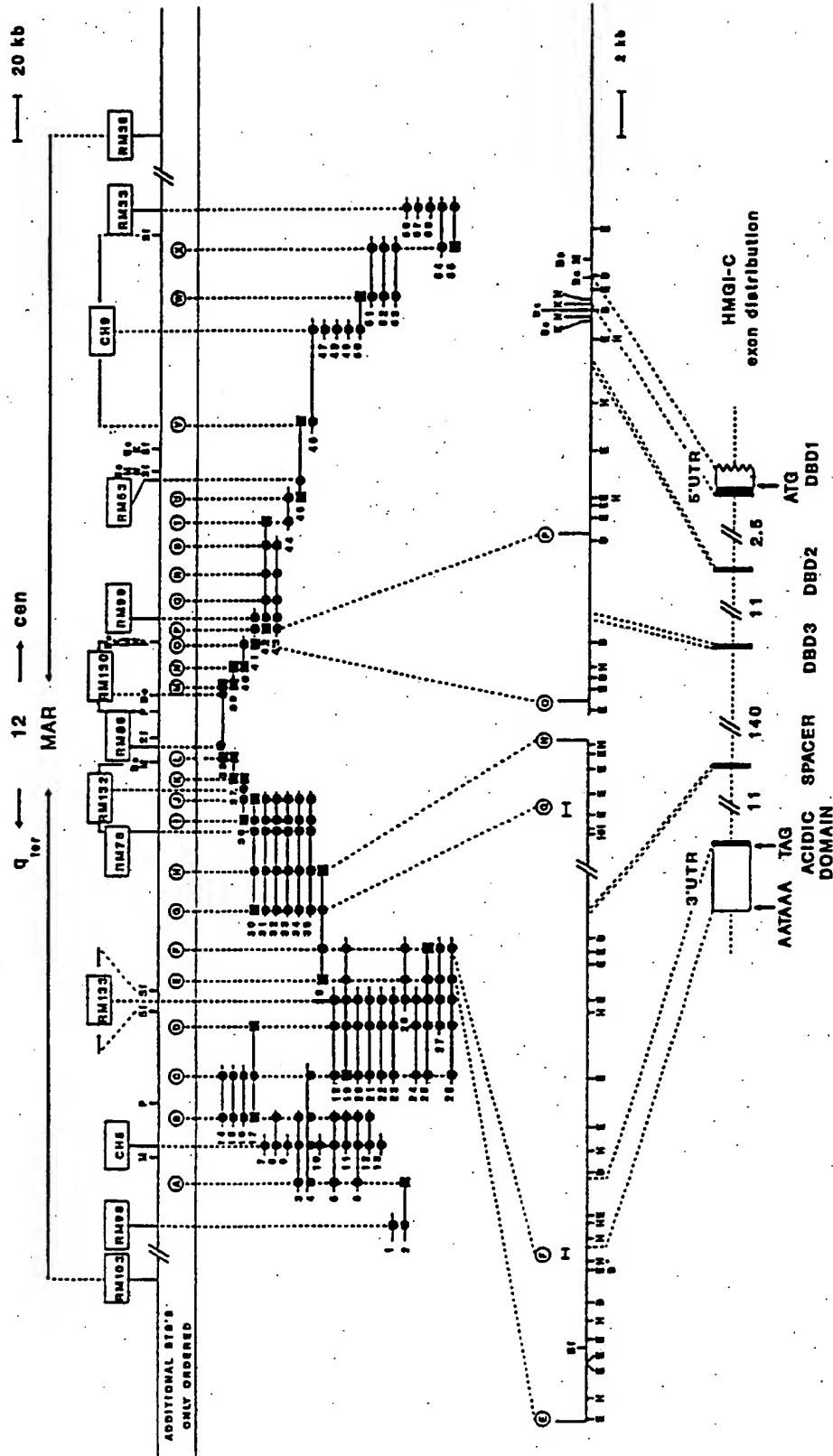


FIG. 2

6/21

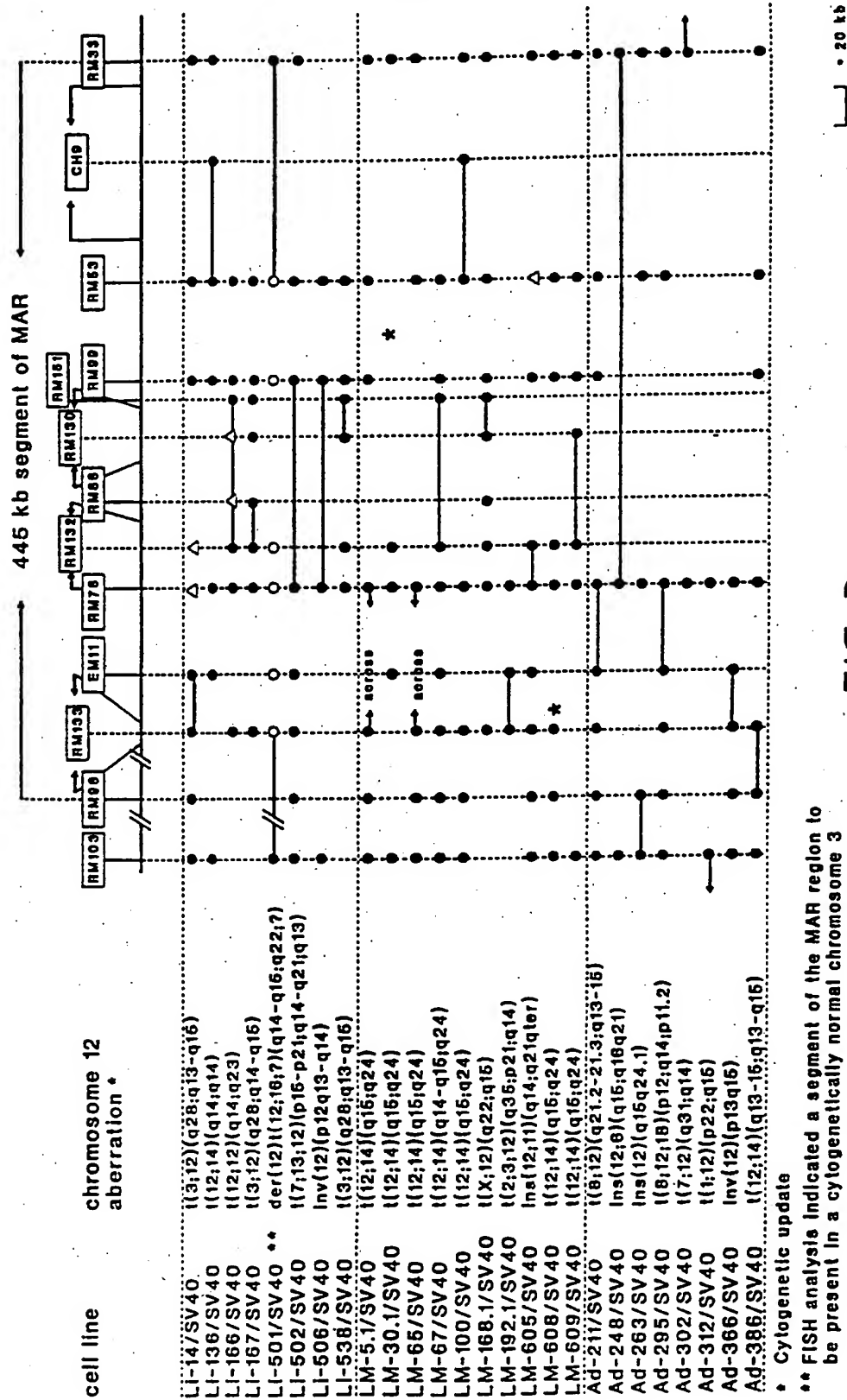


FIG. 3

7/21

FIG. 4

CG CTT CAG AAG AGA GGA <sup>primer</sup> CCG GGC CGC CCC AGG AAG CAG CAG CAA A	45
L Q K R G R G R P R K Q Q Q K	
AA CCA ACC GGT GAG CCC TCT CCT AAG AGA CCC AGG GGA AGA CCC A	90
P T G E P S P K R P R G R P K	
AA GGC AGC AAA AAC AAG AGT CCC TCT AAA GCA GCT CAA GAG GAA G	135
G S K N K S P S K A A Q E E A	
CA GAA GCC ACT GAA GAA AAA CGG CCA AGG GGC AGA CCT AGG AAA T	180
E A T E E K R P R G R P R K W	
MMG <sup>4pp</sup> GGT GGT GGC CAT TCA GGG CAA CTG GGG CCT TCG TCA GTT GCC CCT T	225
G G H S G Q L G P S S V A P S	
CA TTC CGC CCA GAG GAT GAG CTT GAG CAC CTG ACC AAA AAG ATG C	270
F R P E D E L E H L T K K M L	
TG TAT GAC ATG GAA AAT CCA CCT GCT GAC GAA TAC TTT GGC CGC T	315
Y D M E N P P A D E Y F G R C	
GT GCT CGC TGT GGA GAA AAC GTA GTT GGG GAA GGT ACA GGA TGC A	360
A R C G E N V V G E G T G C T	
CT GCC ATG GAT CAG GTC TTC CAC GTG GAT TGT TTT ACC TGC ATC A	405
A M D Q V F H V D C F T C I I	
TC TGC AAC AAC AAG CTC CGA GGG CAG CCA TTC TAT GCT GTG GAA A	450
C N N K L R G Q P F Y A V E K	
AG AAA GCA TAC TGC GAG CCC TGC TAC ATT AAT ACT CTG GAG CAG T	495
K A Y C E P C Y I N T L E Q C	
GC AAT GTG TGT TCC AAG CCC ATC ATG GAG CGG ATT CTC CGA GCC A	540
N V C S K P I M E R I L R A T	
CC GGG AAG GCC TAT CAT CCT CAC TGT TTC ACC TGC GTG ATG TGC C	585
G K A Y H P H C F T C V M C H	
AC CGC AGC CTG GAT GGG ATC CCA TTC ACT GTG GAT GCT GGC GGG C	630
R S L D G I P F T V D A G G L	
TC ATT CAC TGC ATT GAG GAC TTC CAC AAG AAA TTT GCC CCG CGG T	675
I H C I E D F H K K F A P R C	
GT TCT GTG TGC AAG GAG CCT ATT ATG CCA GCC CCG GGC CAG GAG G	720
S V C K E P I M P A P G Q E E	
AG ACT GTC CGT ATT GTG GCT TTG GAT CGA GAT TTC CAT GTT CAC T	765
T V R I V A L D R D F H V H C	
GC TAC CGA TGC GAG GAT TGC GGT GGT CTC CTG TCT GAA GGA GAT A	810
Y R C E D C G G L L S E G D N	
AC CAA GGC TGC TAC CCC TTG GAT GGG CAC ATC CTC TGC AAG ACC T	855
Q G C Y P L D G H I L C K T C	
GC AAC TCT GCC CGC ATC AGG GTG TTG ACC GCC AAG GCG AGC ACT G	900
N S A R I R V L T A K A S T D	
AC CTT TAG ATT CAG TCA CCT GTT CAG CCG GCA CTG AGA AGA ACG A	945
L * I Q S P V Q P A L R R T N	
AC ACA AGA AAA AGA TAA GAA ATA CTA GAG TAA AGG CCA TCA AAC T	990
T R K R * E I L E * R P S N Y	
AC GCG AAA <sup>primer</sup> AAA AAA AAA AAA AAA GAT GTC GAC GGA TCC TT	1033
A K K K K K K K D V D G S :	

8/21

FIG. 5-1"GEN3" → *partial cDNA sequence*

GTCACITTTTA	TTTGGGGGTG	TGGACAGCTG	CTTTCCCAGG	GGAGTACTTC	TTACAGTGGG	60
ATTTCAAGAC	AAGATCGGCC	TGAAGAAAAA	TTATATTTGT	ATATTTTTTA	AAAAGTGGGA	120
ACTTTGAGGC	TCAGAGACAG	AGCAGAAGAC	AGAACCTGGT	CTTCTGATTC	CCTGTGTTCT	180
GCTTTTTTCA	TTGTTCCACT	GGACGCTCAT	CAGAGGGAAG	ATCTTTTTTC	TCAATTGATT	240
CCAACATGT	CTCACCCATC	TTGGCTGCCA	CCCAAAAGCA	CTGGTGAGCC	CCTCGGCCAT	300
GTGCCTGCAC	GGATGGAGAC	CACCCATTCC	TTTGGGAACC	CCAGCATTTC	AGTGTCTACA	360
CAACAGCCAC	CCAAAAAGTT	TGCCCCGGTA	GTTGCTCCAA	AACCTAAGTA	CAACCCATAC	420
AAACAACCTG	GAGGTGAGGG	TGATTTTCTT	CCACCCCCAC	CTCCACCTCT	AGATGATTCC	480
AGTGCCCTTC	CATCTATCTC	TGGAACTTTT	CCTCCTCCAC	CACCTCTTGA	TGAAGAGGCT	540
TTCAAAGTAC	AGGGGAATCC	CGGAGGCAAG	ACACTTGAGG	AGAGGCGCTC	CAGCCTGGAC	600
GCTGAGATTG	ACTCCTTGAC	CAGCATCTTG	GCTGACCTTG	AGTGCAGCTC	CCCCTATAAG	660
CCTCGGCCTC	CACAGAGCTC	CACTGGTTCA	ACAGCCTCTC	CTCCAGTTTC	GACCCCACTC	720
ACAGGACACA	AGAGAATGTT	CATCCCGAAC	CAACCCCTTC	TAACAGCAAC	CAAGAAGTCT	780
ACATTGAAAC	CACAGCCTGC	ACCCCAGGCT	GGACCCATCC	CTGTGCCTCC	AATCGGAACA	840
CTCAAACCCC	AGCCTCAGCC	AGTCCAGGCC	TCCTACACCA	CGGCCTCCAC	TTCTTCAAGG	900
CCTAUCCTTTA	ATGTGCAGGT	GAAGTCAGCC	CAGCCCAGCC	CTCATTATAT	GGCTGCCCGT	960
TCATCAGGAC	AAATTATGTT	CTCAGGGCCC	CAGGGCTATA	ACACTCAGCC	AGTTCCCTGC	1020
CTGGGUCAGT	GTCCACCTCC	TTCAACACGG	GGAGGCATGG	ATTATGCCTA	CATTCCACCA	1080
CCAGGACTTC	AGCCGGAGCC	TGGGTATGGG	TATGCCCCCA	ACCAGGGACG	CTATTATGAA	1140
GGCTACTATG	CAGCAGGGCC	AGGCTATGGG	GGCAGAAATG	ACTCTGACCC	TACCTATGGT	1200
CAACAAGGTC	ACCCAAATAC	CTGGAAACGG	GAACCAGGGT	ACACTCCTCC	TGGAGCAGGG	1260
AACCAGAACC	CTCCTGGGAT	GTATCCAGTC	ACTGGTCCCA	AGAAGACCTA	TATCACAGAT	1320
CCTGTTTCAG	CCCCCTGTGC	GCCACCATTC	CAGCCAAAGG	GTGGCCATTC	AGGGCAACTG	1380
GGGCCTTCGT	CAGTTGCCCC	TTCAATCCGC	CCAGAGGATG	AGCTTGAGCA	CCTGACCAAA	1440
AAGATGCTGT	ATGACATGGA	AAATCCACCT	GCTGACGAAT	ACTTTGGCCG	CTGTGCTCGC	1500
TCTCGAGAAA	ACGTAGTTGG	GGAAGGTACA	GGATGCACTG	CCATGGATCA	GGTCTTCCAC	1560
GTGGATTGTT	TTACCTGCAT	CATCTGCAAC	AACAAGCTCC	GAGGGCAGCC	ATTCTATGCT	1620

9/21

FIG. 5-2

GTGCAAAAGA AAGCATACTG CGAGCCCTGC TACATTAATA CTCTGGAGCA GTGCAATGTG 1680  
 TGTTCCAAGC CCATCATGGA GCGGATTCTC CGAGCCACCG GGAAGGCCTA TCATCCTCAC 1740  
 TGTTCACCT GCGTGATGTG CCACUGCAGC CTGGATGGGA TCCCATTAC TGTGGATGCT 1800  
 GCGGGGCTCA TTCACTGCAT TWAGGACTTC CACAAGAAAT TTGCCCCGCG GTGTTCTGTG 1860  
 TGCAAGGAGC CTATTATGCC AGCCCCGGGC CAGGAGGAGA CTGTCCGTAT TGTGGCTTTG 1920  
 GATCGAGATT TCCATGTTCA CTGCTACCGA TGCGAGGATT GCGGTGGTCT CCTGTCTGAA 1980  
 GGAGATAACC AAGGCTGCTA CCCCTTGGAT GGGCACATCC TCTGCAAGAC CTGCAACTCT 2040  
 GCCCCGATCA GGGTGTTCAC CGCCAAGGCG AGCACTGACC <sup>SEP</sup>TTTAGATTCA GTCACCTGTT 2100  
 CAGCCGGCAC TGAGAAGAAC GAACACAAGA AAAAGATAAG AAATACTAGA GTAAAGGCCA 2160  
 TCAAACCTACG CGATAGTCAC TGTTCCTCAT CTGCTATTAA CCTTGCCTTA GAAACACATA 2220  
 AATTATGAGA TTTTTTTTTA AAAGTTGTTA CCAATACAC ATTTACATT GAATCATGTA 2280  
 GGATCTTGAT GGGCCTTTGT TCCCAAGGAC TTCCACATTT TTGCACAGAT TATGCTCCAT 2340  
 CCCTTCACTT CTGCATTGCT GTAACTTTTA ATCCCTATGT TTGTCTCACT TTTCATCTGG 2400  
 TTGAATGGCT TTTCTTAGTG TGGTATTTGC TGTACATAG TTTTTCTG GGTGAGTCTG 2460  
 CCAACTCACA GGTGCTTTTA GGCTTGAAT CTCCATCCTA TCATTTCCGT TTTGCTGTG 2520  
 ACTGTAAAGA GTAGCCATTC TTTTCCCATG TATTGAAGAG GATATTCTTC TCTTGCTTTA 2580  
 TACTACTCAC GTTCCTGGGG AGGGAAATGC ACAATTTTTT TTTGTTAGGC TGTAAAGAAT 2640  
 TTAAGCTGTA AATTACATAA GTTAGAACAA GCCCAAATTT AATTGCAAC CATCAGAATT 2700  
 CAGAATCTAT AGTGACCACT GATCAAGGCT AATTGGAAAA GAGTTATCGG CCCATAGCTA 2760  
 ATAAGTAGTG ACAGACAACC AAGCTTCAAT ATTTTTCTAA AGAAATTACA GGTGGGATAT 2820  
 GCTAGAAAAG GCATTTTGGG GTTATGTTTA AAAAAACATT ATTGTCCAC AATATTACCT 2880  
 TAAGATTTTT CTTTCCGCA CTACCTGAAC ATTGTAATAC AGACAACTT GATTCTTCT 2940  
 AGAAGATAAC ATTTTCAATA CTGTCCCACT TGTATCTTA AAAATATTGT CATGTTTATT 3000  
 CTAATATCCA ACGCAACTAT CAAAATGCC TTTTCTCTA GAGGATGAAG GCTGTGAAAA 3060  
 AACCGTTCAA ATTCTCTCT TTTTCTTTTT TATTACCAGG TCCATTTTGC CTGACAATTG 3120  
 CAAATCAGAG CATACAAAAT AAAACTGTGC AGTTTGTGTT GGTTTACTTT CAAAAGAGTA 3180  
 GAAAGCTTGA AAAGATTCTG AAACCACAGT TTCAATTATC TCATAATCCT TCTGCAACTG 3240  
 AATTACATA TTGCAGGAGA CATTTTCATA TCATCAATGT GACATTTACA CCACACTTTC 3300  
 AAAGACAATC ACTGAAACAA AAATTGTCTT TATCAGCTAA AAATATGCAG AATCTCTGCC 3360

10/21

FIG. 5-3

TAGAATCTTT	ATTCAAACCT	TTATTAGCCA	GTGAAACACT	TGCTTGCCAA	CTGCCAAGCC	3420
ATACTTATTA	AGTTCGAACA	TGTTTCACTT	AAGGAGAGAC	ACCTAGCTTA	GTCATGGCAA	3480
GTTGCCATTT	TGTAAACTAA	GGATTTTGGG	CTGAGATTTT	TTAAATCTTT	CTTCAAATCT	3540
CCCACAAGTA	TATACTTTTA	AATTATGGAG	TATTTTAAGT	CTACAAAAG	GTATAAATAA	3600
TAATATAATG	AATTCCTATA	TACCTAATAC	CCAGTTTAAG	ACACCAAATA	TAACAAGTAT	3660
AATTACATCC	TCCAATGTAC	CGTTTCCTTA	TTCCACAGAT	ATCTTTTTC	TTATTGTGAA	3720
GTGATGTTCA	GATTTCTAGT	TTTTTTTTCT	AGTTTTTAAT	TTTAACATCA	GAAGTGAAT	3780
AAAAAATTAT	GGATACGTGT	TTTGAATTGC	AACTATTCC	TCAGGAATTC	CAATTAAATT	3840
TATTTTACTT	GAATAGGAAT	GATCATAAAA	GTGATTCTTT	TTTTGTGACT	AGAAATCTTT	3900
AAGCCGATGG	TCACTATAGC	TCATCCTTAA	TGTATGGCTC	ATTTGCTTTT	GTCACTAAAC	3960
GGTTTTGTGT	TAGAACCACC	AAAATTATAG	CTTTTAAGAG	CTTCCTTTGA	CCACTGTCTT	4020
TTTCTTACCC	TACTTCTCTT	ATCTTTGATC	GTATATTCTT	CATAAATGTA	AAATATGATGA	4080
GATTCACCTA	GGGGCAGCAT	GTTAGTTTTC	GGAGGCAATG	TCAACTGTGT	CTCTGAATTC	4140
CTGTCTTCCA	AATTGAAGCC	AGACCATGCT	GATGACCTCA	AGTAGCACTG	ACTATTTGAC	4200
AATAGGGCTG	ATAATGTAA	CGGCATGAAT	TTTGACTTAG	TAACITTTTA	TGTAATACTT	4260
TCGGAGAAAT	TCTCTTAGG	ACAAAGCAGA	GAGTCCAAAT	TATTGAGGGA	TAGATTGTAT	4320
CTC						4323

11/21

FIG. 6

MSHPSWLPPKSTGEPLGHVPARMETTESFGNPSISVSTQOPPARKFAPVVAPEPKY  
 NPYKQPGGEGDFLPPPPPLDDSSALPSISGNFPPPPPLDEEAFKVQGNPGGKIL  
 EERRSSLDAEIDSLTSLADLECSSPYKPRPQSSSTGSTASPFVSTPVTGHEKRV  
 IPNQPLTATKKSTLKQPAPQAGPIPVAPIGTLKPOQPVPAASYTTASTSSRPT  
 FNVQVKSQAQSPHYMAAPSSGOIYGSGPOGYINTQPVFVSGQCFPPSTRGGMDYAY  
 IPPPGLQPEPGYGYAPNQGRYYEGYYAAGPGYGGENDSDPTYGQOQHPNTWKREP  
 GYTTPGAGNQNPFGMPVTGPKRTYITDFVSAPCAPFLQPKGGHSGQLGPSSVAP  
 SFRPEDELEHLTKOMLYDMENPPADEYF ERCARCGENVVGEGTGCTAMDQVFHVD  
 CFTCLCNKLRGQPFYAVEKKAYCEPCYINTLE QCNCVCSKPIMERILRATGKAY  
 HPFCFTCMCHESLDGIPFTVDAGGLTHCIEDF KKFAPRCVCKEPIMPAPGQE  
 ETVRIVALDDEDFHVHCYRCEDCGGLISEGDNQGCYPLDGHILCKTCNSARIEVLT  
 AKASTDL\*

12/21

FIG. 7-1

 BASE COUNT 1119 a 988 c 791 g 1169 t  
 ORIGIN

```

1  ottgaatctt  ggggcaggaa  ctoagaaaac  ttccagcccg  ggcagccggc  gcttggtgca
61  agaotoagga  gctagagagc  gtoocccctc  cgaotctccg  gtgocgccgc  tgcctgctoc
121  agooaacola  gggaggaggg  tgoaacccac  laototgtcc  tctgcctgtg  ctccgtgccc
181  gaootatcco  ogggggagtc  looccatcct  cotllgcttt  cggactgccc  aaggaoattt
241  oaatotoaat  ctctctcttc  loletotolo  totctctgtc  tctctctctc  tctctctctc
301  totulutoto  gcagggtggg  gggagagaga  gggagaaatc  ttcccccggc  taaoatttca
361  aggaumumaa  ttoactocaa  gtctcttccc  lllccaaagg  gcttccgaag  tgctocoggt
421  gocogoaaci  uolgatocaa  accogggaga  gggagctctg  cgaacctaaa  gctctctctc
481  ottctccctc  gotttccctc  tootottgct  aottccacct  coaacggcac  ctocaoctoc
541  ggcactcaao  oacogggcgc  gccgcacccg  gcagcgcttc  ctctctctct  ccttctcttc
601  ootctctctc  tttttggcag  ccggtggcag  tccggtgttg  atggtggcag  ogggggcagc
661  ctatgcacaa  gcagccctcg  cagcccgcaa  gctcgccttc  gccocggcgg  cgtocccagc
721  cctatucotc  catctcccca  aaggltgtgg  ggaotccgg  ggcggtggag  gccaaggggc
781  ootggagaga  gtagggggcg  gggaggagcg  gtagagcgca  cgcggtgagg  ggcggggcca
841  ggcgtccact  toagccaggg  gaaaaoctgc  cggccagcgg  cctcagaaga  gaggacggcg
901  oocgococag  aayuaagaga  aagaacocaa  cgggtgagcc  tctcctaaga  gacctagggg
961  aagaccocaa  ggaagocaaa  acaagagluc  ctctaaagca  gctcaaaaga  aagcagaagc
1021  ootggagaga  aaggggocaa  gaggcagacc  laggaatgg  ocaacacag  ttgttcagaa
1081  gaagootgct  caggagggaa  ctgaagagac  atootacaa  gagtctgccg  aagaggacta
1141  gggggggcaa  cgttggallt  ctacotcagc  agcagttgga  tcttttgagg  ggaagagaca
1201  ctgcagtgac  caollattot  gtattgocat  ggtctttoa  ctctctctct  ggttgggggt
1261  ggttgggggt  ggggaggggg  ggttgggggt  gggagaaatc  acataaoctt  aaaaaggact
1321  atattaatca  ootctctctc  aatcccttca  cagtcocagg  tttagtgaag  aactgtgtga
1381  aaoacagggg  aacacagctc  acaatggaaa  ttttaattac  tgttttcttt  ttctttaaacc
1441  taataatagt  ttgttgatct  gataagcaag  agtggggcgg  tgayaaaaac  cgaattgggt
1501  ttagtgaagg  aclycaotgc  atgcaacaca  gaaacgtgto  aacttttgta  cgtcgggcat
1561  tcatatagga  agaacggcgt  gtgtaaocct  gtgtaaocct  oaaataocac  coaacccac
1621  toctctgagt  gaatocicly  tttagaaoao  caaagataag  gactagatag  tactttctct
1681  ttttggtata  atottglaya  oacttaottg  atgttttta  actttttatt  tctaaatgag
1741  acaagagggg  gatgtatoot  ttcattcaga  taacaaacta  gaaaaggtta  tgttaatttt
1801  tcaaaaaggg  aagtaagcaa  aaaaattttg  oaaactcttc  tatttatgga  tatcacacat
1861  atcagcagga  gtaataaati  laotoacaga  acttgttttc  aggaacaaac  ttoattttca
1921  ggaallotac  ttctacaga  gaoaaaatgc  cattagcaa  taataaacac  ttgtocaggt
1981  oagagcaatt  aaggaacata  gaaaggttaa  attatootot  ttgttaattt  algaagaggt
2041  aaaaagagat  aatgcattgt  gaactocool  aattatgagg  tgggaggago  gaaatotaaa
2101  tttcttttgc  tatagttola  ootcaattta  aaaaagcaaa  aaaaaagggg  gggggcaato
2161  ttctctctgt  tctttctctc  tctctctctc  tctctctctc  tctttctatg  tgtatcagtt
2221  tctatgaaag  aactgaatac  cacllaucic  aaatlaagca  tatgtgttac  ttcaagttat
2281  aacttttgac  ataagatggt  tgacnaaggt  gcttttcttc  quettgagtt  caccatctct
2341  toatloaaac  tgcactttta  gccagagatg  caatatatcc  coactactoa  atactacotc
2401  tgaatgttac  aagaaattta  cagtctagta  ootattacal  gctgctatcc  acaagcaatg
2461  oaaagaaaaa  acttaotggg  taggtgatto  taatcatotg  cagttctttt  tgtacactta
2521  attacagtta  aagaaagcat  ctctctctct  tgtttoagca  tgactatgta  tttttotatg
2581  ltttttttat  taanaatttt  taanaacttt  gtttoagctt  otctgctaga  ttctacatt
2641  aacttgaana  ttttttaacc  aactcgctcc  taggtcttta  aggaataatt  tctctaatca
2701  caatacaaat  oacacaaagt  ttgactgtaa  tatttaata  ttacootcca  agtctgtacc
2761  toaaatgast  totttaagga  yatggactaa  ttgacttgca  aagacotacc  tccagacttc
2821  aaaaaggaatg  aacttgttac  ttgcagcatt  catttgtttt  ttoaatgttt  gaaatagtto
2881  aaoctgagca  taocclagt  oaaaactatt  ttgttaaaag  aactttgata  gaaaaggaag
2941  cgttttttca  taottttgca  aaataagtaa  aaataaata  aaataagcc  aaoottaaaa
3001  gaoottgag  otttgtaggt  gagatgcaac  aagocctggt  ttggaatant  gaaatcaaaa
3061  atatgtgttt  ttaagattag  ttgaatataa  gaaaatgctt  gaaatatttt  ttatgtattt
3121  ttacacaaat  gtgatttttg  taalutgtct  caacagattt  tattttaaac  gcltcttatg
3181  tagagttttt  atgocctttc  ctctagtgca  gtgtgctgac  ttttaacat  ggtattatca
3241  actggggcag  gaggtagttt  ctcatgucgg  cttttgtcag  tatggctttt  agtactgaag
3301  ccaaatgaaa  ctcaaaacca  tctctcttcc  agotgottoc  gggaggtagt  ttoaaagggc
3361  acatacctct  ctgagactgg  cagatogcic  actgttgtga  atoccaaag  gagctatgga
3421  gaggattaaa  actcaagatt  actgttaact  gtgggttaaa  taagcaata  aaoagtgggt
3481  oataaaaata  aaagtggcat  tncatatctt  tggatggggc  ttttgaanaa  ctcatggccc
3541  agctaataaa  atggaaycaa  ttgotcatgt  tgggocaaaa  tggtgocagg  agtgatttcc
3601  atctcttgta  aagttacact  ttattttctc  gtagtgtgta  caataaaaac  acactaciau
3661  ctcttaagtc  ccagtatacc  tcaillitaa  taotgaaaaa  aaaaagcttg  ggcacatgga

```

STOP

STOP

**FIG. 7-2**

3721 ccagtaagaa uuuuataaaa ttttatata tatagtttat ttttctggaa gataaatttt  
3781 ataggactgt ttttctgtg ttttggctgc agctacataa gactggacat ttaactttto  
3841 taacatttct gcaagttayy tatgtttgca ggagaaaagt atcaagacgt ttaactggag  
3901 ttgaotttct cactgttctt ttgagtgtct tctaaottta tctttgtto tttaaglaaa  
3961 attgctgtct atgattgtac tttgaaatgc ttgcttcttg aaatattto totagtgtat  
4021 tatcactgtc tttctgcac aataaacata acagcctctg tgatccc

14/21

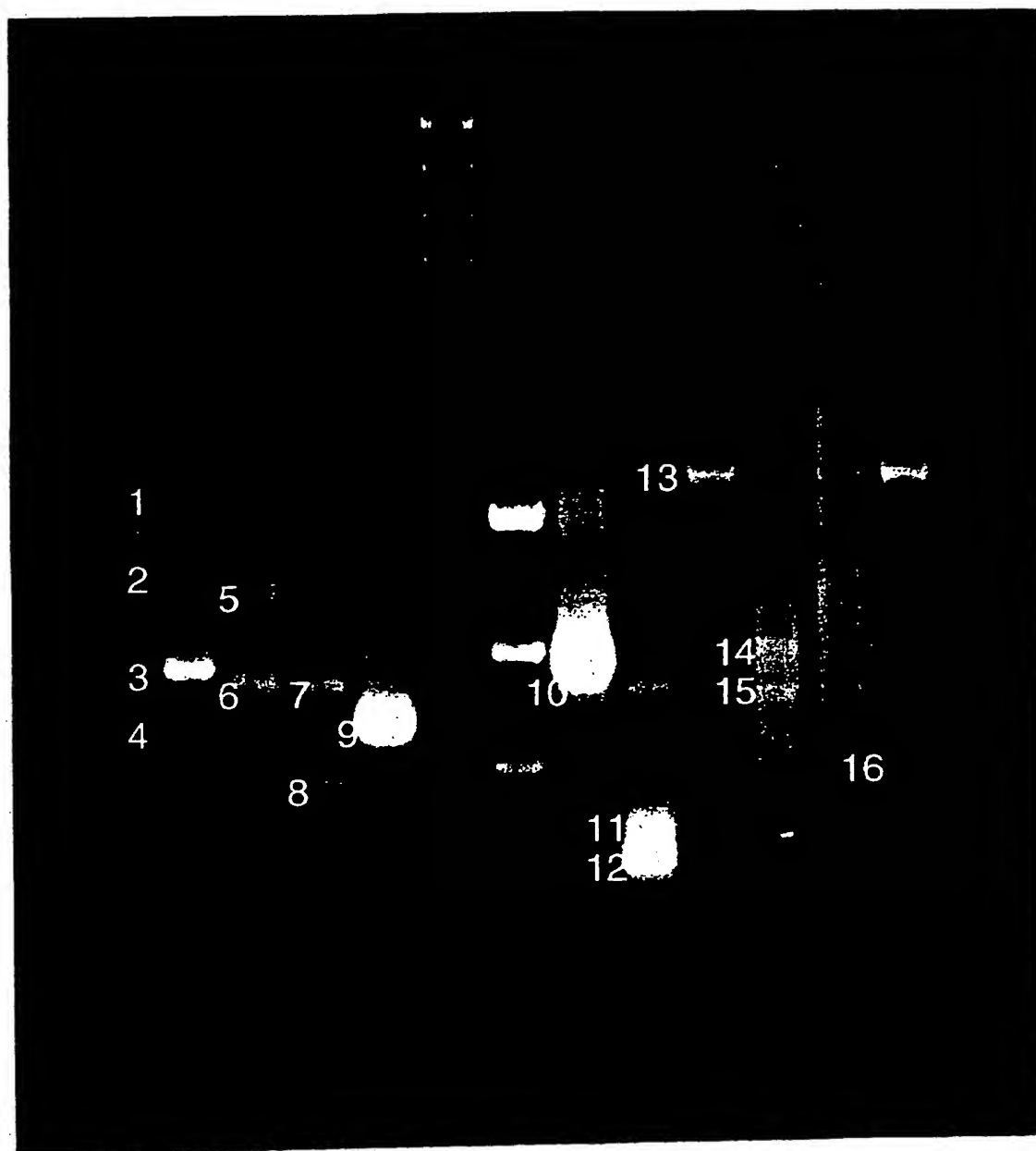


FIG. 8

Figure 1 to Annex 1

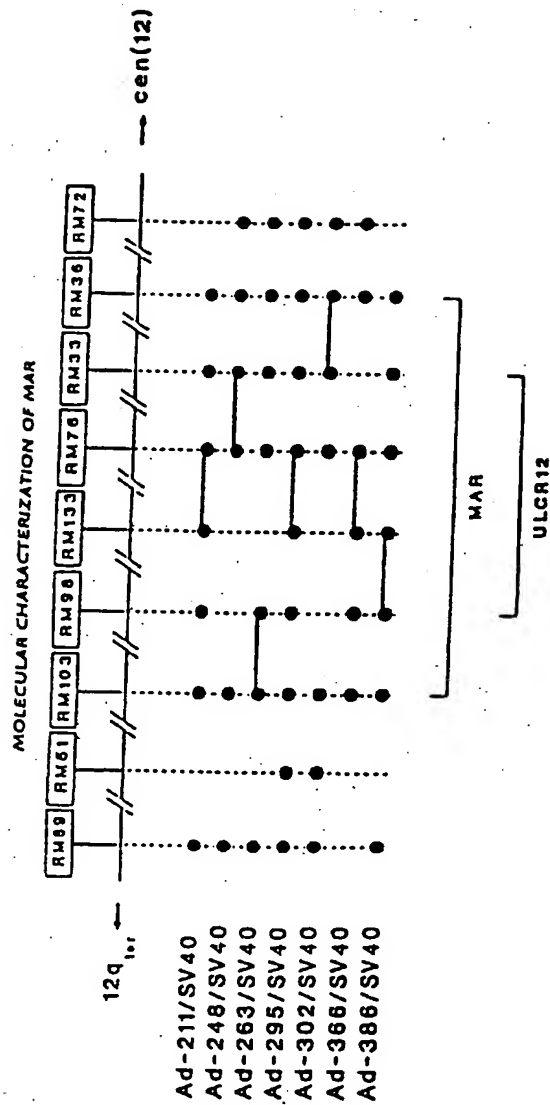
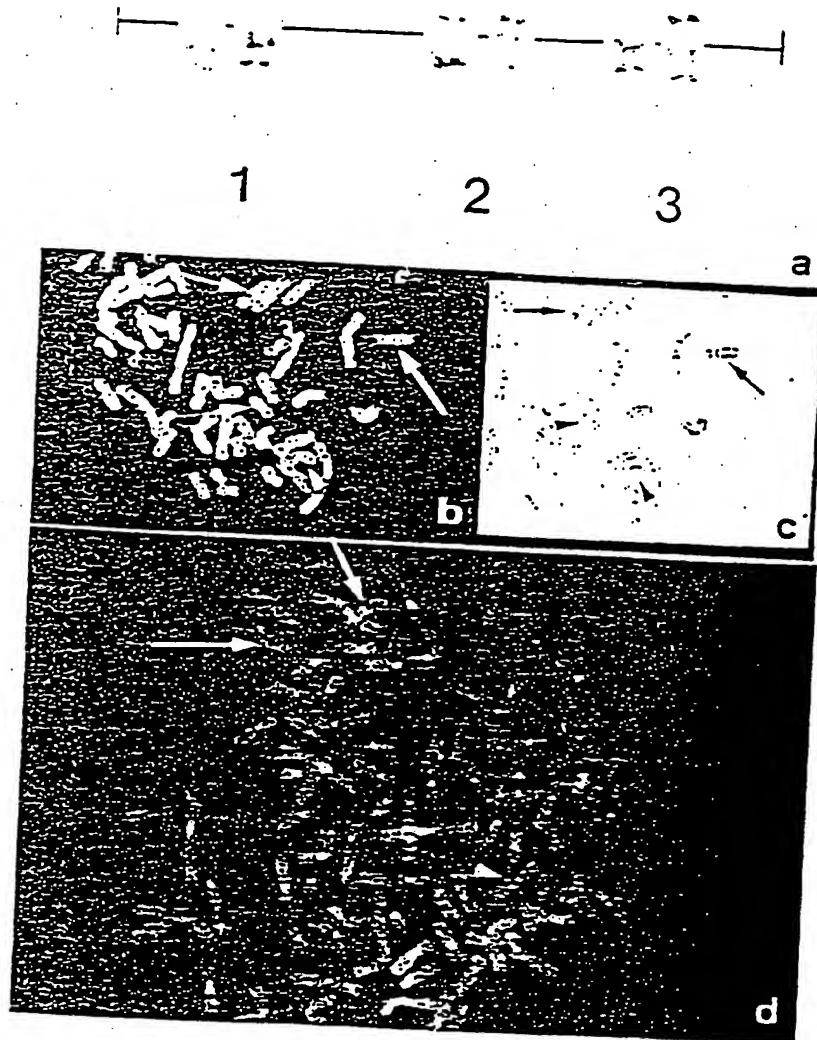


Figure 2 to Annex 1



17/21

Figure 3 to Annex 1

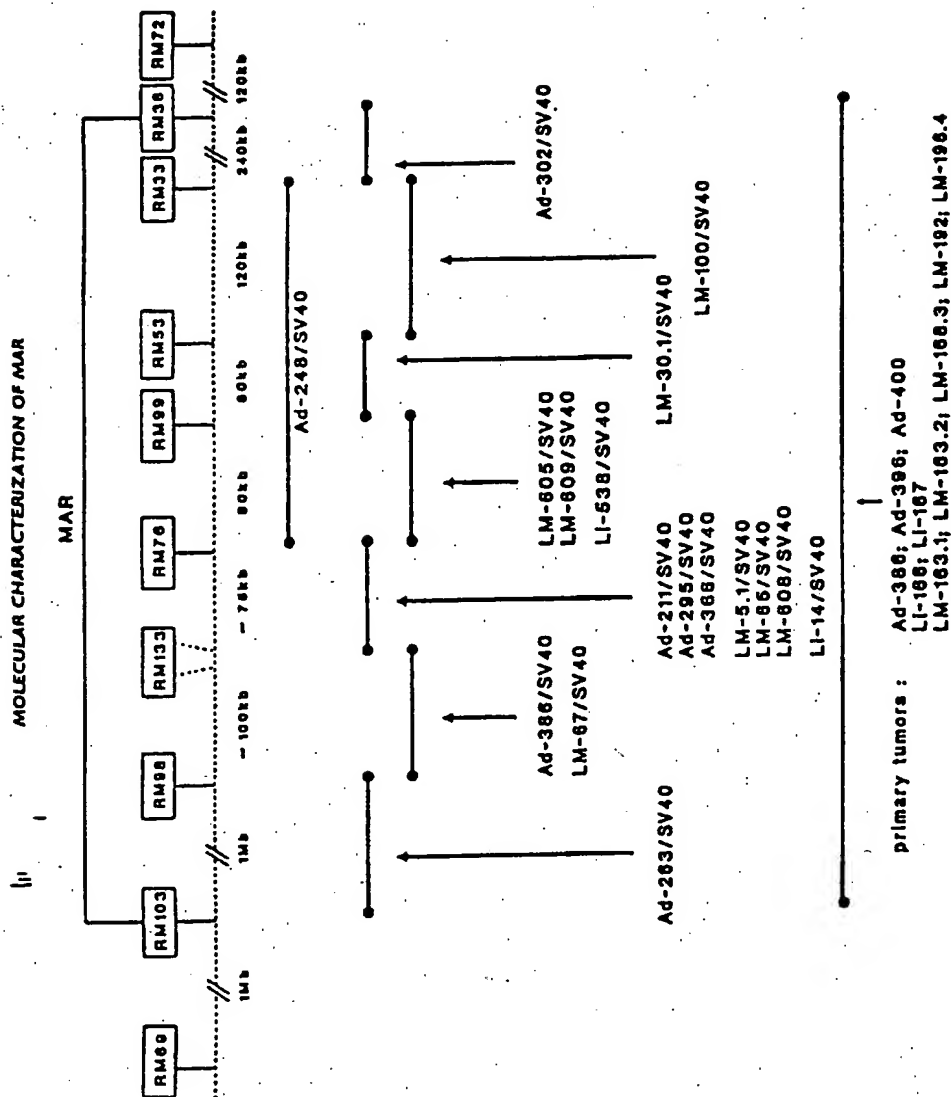
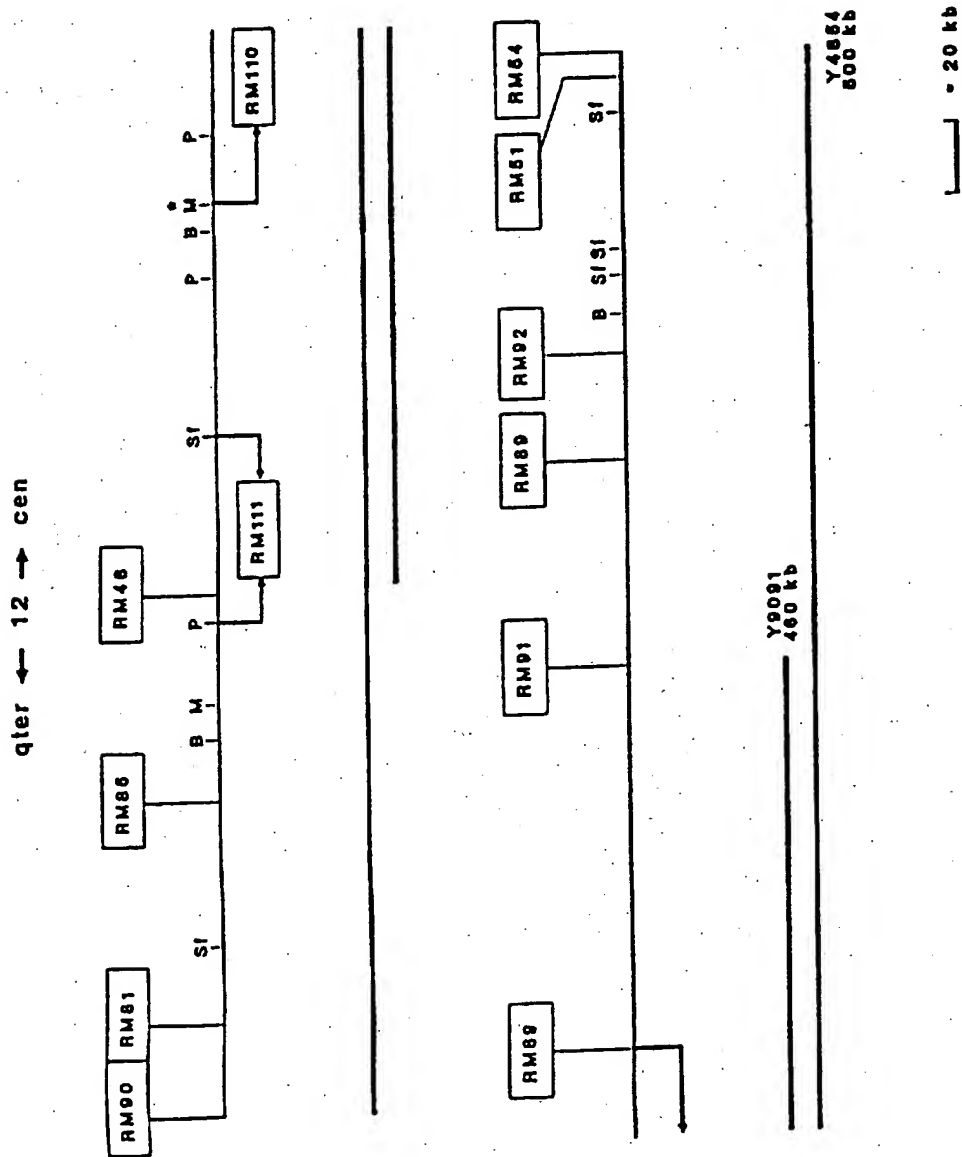


Figure 1 of Annex 2



19/21

Figure 2 of Annex 2

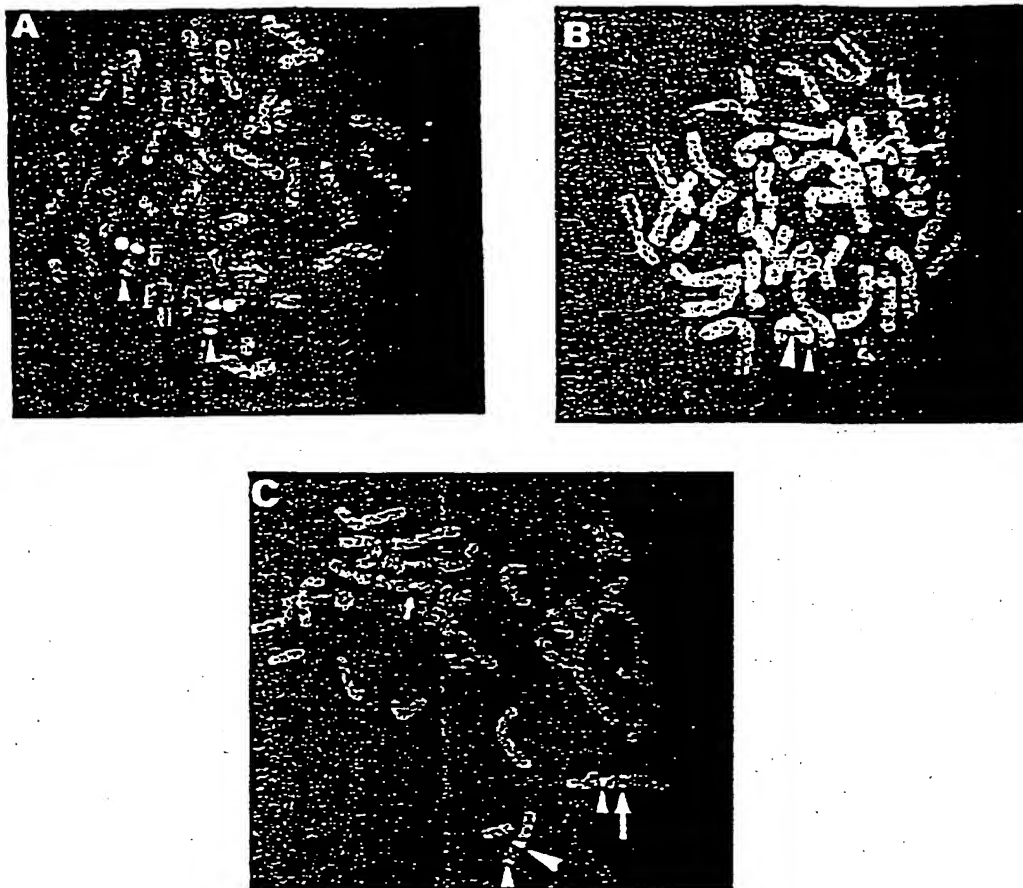


Figure 3 of Annex 2

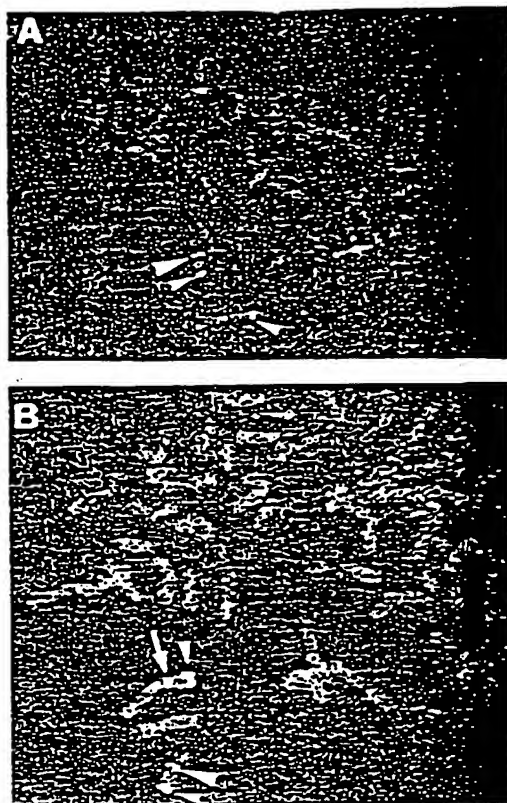
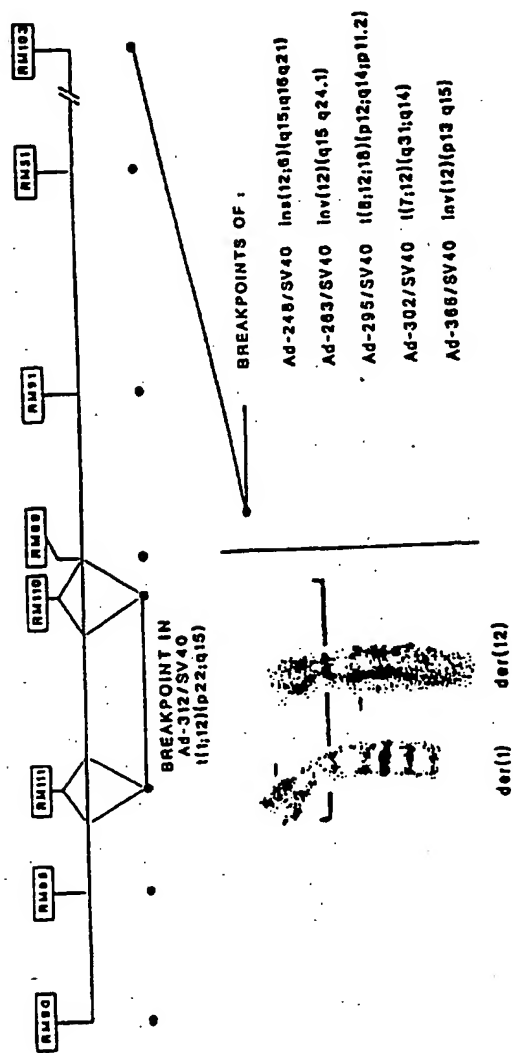


Figure 4 of Annex 2



# INTERNATIONAL SEARCH REPORT

Inter-Application No  
PCT/EP 96/00716

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N15/11 C07K14/47 C12Q1/68 C07K19/00  
C07K16/18 A61K38/17 A01K67/027

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 C07K C12Q A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CHEMICAL ABSTRACTS, vol. 122, no. 7, 13 February 1995 Columbus, Ohio, US; abstract no. 73586t, SCHOENMAKERS, ERIC ET AL.: "Identification, molecular cloning and characterization of the chromosome 12 breakpoint cluster region of uterine leiomyomas." page 267; XP002005567 see abstract & GENES, CHROMOSOMES CANCER , vol. 11, no. 2, 1994, pages 106-118, --- -/-	1,2,4-7

☒ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

### \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search

13 June 1996

Date of mailing of the international search report

12.07.96

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+ 31-70) 340-3016

Authorized officer

Delanghe, L

# INTERNATIONAL SEARCH REPORT

Inter Application No  
PCT/LP 96/00716

## C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No. .
X	<p>NUCLEIC ACIDS RESEARCH, vol. 19, no. 24, 1991, OXFORD GB, pages 6793-6797, XP002005564 G.MANFIOLETTI ET AL.: "cDNA cloning of the HMGI-C phosphoprotein ,a nuclear protein associated with neoplastic and undifferentiated phenotypes." see the whole document</p> <p style="text-align: center;">---</p>	1,2,4-7
X	<p>BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 201, no. 1, 30 May 1994, ORLANDO, FL US, pages 63-70, XP002005565 UMESH A. PATEL ET AL.: "Expression and cDNA cloning of human HMGI-C phosphoprotein." see the whole document</p> <p style="text-align: center;">---</p>	1,2,4-7
X	<p>DATABASE WPI Section Ch, Week 8735 Derwent Publications Ltd., London, GB; Class B04, AN 87-245773 XP002005570 &amp; JP,A,62 166 897 ( TOYO SODA MFG KK) , 23 July 1987 see abstract</p> <p style="text-align: center;">---</p>	6
P,X	<p>NUCLEIC ACIDS RESEARCH, vol. 23, no. 21, 1995, OXFORD GB, pages 4262-4266, XP002005566 KAI-YIN CHAU ET AL.: "The gene for the human architectural transcription factor HMGI-C consists of five exons each coding for a distinct functional element." see the whole document</p> <p style="text-align: center;">---</p>	1,2,4-7
P,X	<p>CHEMICAL ABSTRACTS, vol. 123, no. 1, 3 July 1995 Columbus, Ohio, US; abstract no. 6842f, VAN DE VEN,WIM J.M. ET AL.: "Molecular characterization of MAR,a multiple aberration region on human chromosome segment 12q13-q15 implicated in various solid tumors." page 721; XP002005568 see abstract &amp; GENES,CHROMOSOMES CANCER, vol. 12, no. 4, 1995, pages 296-303,</p> <p style="text-align: center;">---</p> <p style="text-align: center;">-/--</p>	1,2,4-7

# INTERNATIONAL SEARCH REPORT

Inter-Application No

PC1/CP 96/00716

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	<p>CHEMICAL ABSTRACTS, vol. 123, no. 15,            9 October 1995            Columbus, Ohio, US;            abstract no. 195315v,            SCHOENMAKERS, ERIC ET AL.: "Recurrent            rearrangements in the high mobility group            protein gene, HMGI-C, in benign mesenchymal            tumors."            page 885;            XP002005569            see abstract            &amp; NAT.GENET.,            vol. 10, no. 4, 1995,            pages 436-444,</p> <p>-----</p>	1

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☒ **BLACK BORDERS**

☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☐ **FADED TEXT OR DRAWING**

☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☒ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**